CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

RECOMENDAÇÃO DE MODELOS DE NEGOCIAÇÃO PARA AÇÕES COM TOMADA DE DECISÃO INTERATIVA MULTICRITÉRIO

RODRIGO MAICON DE ASSIS SILVA

Orientador: Prof. Alisson Marques da Silva
CEFET-MG

BELO HORIZONTE
JULHO DE 2026.

Silva, Rodrigo Maicon de Assis
S586r Recomendação de modelos de negociação para ações com tomada de decisão interativa multicritério / Rodrigo Maicon de Assis Silva. – 2026.
187 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientador: Alisson Marques da Silva.
Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Ações (Finanças) – Teses. 2. Negociação – Teses. 3. Mercado de ações – Previsão – Brasil – Teses. 4. Sistemas de recomendação (filtragem de informações) – Teses. 5. Tomada de decisão multicritério – Teses. 6. Mineração de dados – Teses. 7. Carteiras (Finanças) – Administração – Teses. I. Silva, Alisson Marques da. II. Centro Federal de Educação Tecnológica de Minas Gerais. III. Título.

CDD 530.13

Elaboração da ficha catalográfica pela bibliotecária Jane Marangon Duarte, CRB 6ª 1592 / Cefet/MG

SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO CURSO DE MESTRADO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

"RECOMENDAÇÃO DE MODELOS DE NEGOCIAÇÃO PARA AÇÕES COM TOMADA DE DECISÃO INTERATIVA MULTICRITÉRIO"

Dissertação de Mestrado apresentada por **Rodrigo Maicon de Assis Silva**, em 10 de julho de 2026, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Documento assinado digitalmente
gov.br ALISSON MARQUES DA SILVA
Data: 17/07/2026 13:42:23-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Alisson Marques da Silva (Orientador)
Centro Federal de Educacão Tecnológica de Minas Gerais

Documento assinado digitalmente
gov.br LEANDRO DOS SANTOS MACIEL
Data: 17/07/2026 13:49:58-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Leandro dos Santos Maciel
Universidade de São Paulo

Documento assinado digitalmente
gov.br UAJARA PESSOA ARAUJO
Data: 20/07/2026 11:56:53-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Uajara Pessoa Araujo
Centro Federal de Educação Tecnológica de Minas Gerais

Documento assinado digitalmente
gov.br ARTHUR RODRIGO BOSCO DE MAGALHAES
Data: 20/07/2026 18:10:52-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Arthur Rodrigo Bosco de Magalhães
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof. Dr. Alisson Marques da Silva
Coordenador do Programa de Pós-Graduação em
Modelagem Matemática e Computacional

Agradecimentos

Gostaria de agradecer a todos que participaram da minha vida durante esses dois anos de mestrado. Ao meu orientador, Prof. Alisson, pela confiança desde o processo seletivo; aos amigos que fiz durante o curso, em especial a Rodrigo Coimbra; aos meus pais, Aparecida e Donizete; e à minha melhor amiga, Mariana Sampaio, por sua eterna amizade. Estendo meus agradecimentos à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela bolsa concedida, pois sem ela não teria sido possível concluir esta etapa.

“As pessoas simplesmente detestam a ideia de perder. Qualquer perda, mesmo que pequena, é terrível de se contemplar.” (Daniel Kahneman)

Resumo

A dinâmica do mercado acionário demonstra que ativos de um mesmo setor tendem a apresentar comportamentos similares, ao passo que ações de setores distintos exibem movimentos frequentemente descorrelacionados, o que torna inadequada a adoção de um único modelo de negociação para todo o mercado. Neste trabalho, propõe-se uma abordagem de recomendação multicritério de modelos de negociação para o mercado brasileiro (B3), fundamentada no método TODIM (*TOmada de Decisão Interativa Multicritério*) e na Teoria do Prospecto, com avaliação em dois níveis de granularidade: setorial e individual por ação, sendo este último a principal contribuição do trabalho. O conjunto de dados compreende 239 ações organizadas em oito setores econômicos. Indicadores técnicos são extraídos como variáveis de entrada e, adicionalmente, aplica-se a seleção de atributos com base na Informação Mútua para identificar as *features* mais relevantes, mitigando a presença de variáveis redundantes ou irrelevantes. Doze algoritmos de aprendizado de máquina são treinados por meio de análise *walk-forward* para a geração de sinais de negociação, e seus desempenhos são avaliados com base em oito indicadores de *backtesting*. O algoritmo de recomendação resultante é avaliado em suas versões setorial e individual. Os experimentos, conduzidos no período de alta volatilidade da pandemia de COVID-19 (jan.–jun. 2020), demonstram que ambas as versões superam o Ibovespa em métricas de proteção de capital e retorno ajustado ao risco. Além disso, métricas de generalização testadas confirmam a robustez do algoritmo de recomendação. Por fim, portfólios *long-short* construídos com base nas recomendações demonstraram capacidade significativa de proteção contra o risco sistêmico.

Palavras-chave: Negociação de ações; B3; algoritmo de recomendação; TODIM; aprendizado de máquina; seleção de atributos; tomada de decisão multicritério; portfólio *long-short*.

Abstract

Stock market dynamics show that assets within the same sector tend to exhibit similar behavior. In contrast, stocks from different sectors frequently display uncorrelated movements, making the adoption of a single trading model for the entire market inadequate. In this work, a multicriteria recommendation approach for trading models is proposed for the Brazilian stock market (B3), grounded in the TODIM (*TOmada de Decisão Interativa Multicritério*) method and Prospect Theory, with evaluation at two levels of granularity: sector-level and individual stock-level, the latter being the main contribution of this work. The dataset comprises 239 stocks organized into eight economic sectors. Technical indicators are extracted as input variables, and Mutual Information-based feature selection is applied to identify the most relevant features, thereby mitigating redundancy and irrelevance. Twelve machine learning algorithms are trained via walk-forward analysis to generate trading signals, and their performance is evaluated using eight backtesting indicators. The resulting recommendation algorithm is assessed in both its sector-level and individual stock-level versions. Experiments conducted during the high-volatility period of the COVID-19 pandemic (Jan.–Jun. 2020) demonstrate that both versions outperform the Ibovespa index in capital protection and risk-adjusted return metrics. Furthermore, generalization metrics confirm the robustness of the recommendation algorithm. Finally, long-short portfolios constructed using the recommendations demonstrated a significant capacity to protect against systemic risk.

Keywords: Stock trading; B3; recommendation algorithm; TODIM; machine learning; feature selection; multi-criteria decision-making; long-short portfolio.

Lista de Figuras

Lista de Tabelas

Lista de Algoritmos

Lista de Abreviaturas e Siglas

3WD	*Three-way Decision*
ADF	*Augmented Dickey-Fuller*
AD	*Accumulation/Distribution Line*
ADX	*Average Directional Index*
AE	*Autoencoder*
AELSTM	*Automatic Encoder LSTM*
AES	*Adaptive Exponential Smoothing*
AIC	*Akaike Information Criterion*
ANFIS	*Adaptive Neural Fuzzy Inference System*
ANN	*Artificial Neural Network*
AO	*Aquila Optimizer*
ARIMA	*Autoregressive Integrated Moving Average*
ARR	*Annualized Return Rate*
ART	*Average Recovery Time*
ART²	*Average True Range*
ASR	*Annualized Sharpe Ratio*
ASRWP	*ASR Weighted Portfolio*
B3	Brasil, Bolsa, Balcão
B3ICS	*B3 Index Constituent Stocks*
BA	*Bat Algorithm*
BAH	*Buy-and-Hold*
BandPer	*Bollinger %b*
BandWid	*Bollinger Bandwidth*
BIC	*Bayesian Information Criterion*

BiGRU	*Bi-directional Gated Recurrent Unit*
BiLSTM	*Bi-directional Long Short-Term Memory*
BM	*Basic Materials* (Setor de Materiais Básicos)
BP	*Back Propagation*
BPNN	*Back-Propagation Neural Network*
BPTT	*Back-propagation Through Time*
CAR	*Calmar Ratio*
CART	*Classification and Regression Trees*
CC	*Consumer Cyclical* (Setor de Bens de Consumo Cíclicos)
CCI	*Commodity Channel Index*
CG	*Cumulative Gain*
CNC	*Consumer Non Cyclical* (Setor de Consumo Não Cíclico)
CNN	*Convolutional Neural Network*
CMF	*Chaikin Money Flow*
CO	*Chaikin Oscillator*
COM	*Communication* (Setor de Comunicação)
CPT	*Cumulative Prospect Theory* (Teoria do Prospecto Cumulativa)
CR	*Coverage Rate*
DBN	*Deep Belief Network*
DCG	*Discounted Cumulative Gain*
DIS	*Disparity*
DQN	*Deep Q-Network*
DT	*Decision Tree*
EFS	*Evolving Fuzzy Systems* (Sistemas Fuzzy Evolutivos)
EGARCH	*Exponential Generalized Autoregressive Conditional Heteroskedasticity*
eGNN	*Evolving Granular Neural Network*

ELSTM	*Embedded Layer LSTM*
EMA	*Exponential Moving Average*
eMG	*Multivariable Gaussian Evolving Fuzzy Modeling System*
eOGS	*Evolving Optimal Granular System*
EOM	*Ease of Movement*
ESG	*Environmental, Social and Governance*
ETF	*Exchange-Traded Fund*
EWP	*Equal Weight Portfolio*
FI	*Force Index*
FIN	*Financial* (Setor Financeiro)
FNN	*Feedforward Neural Network*
GA	*Genetic Algorithm*
GAN	*Generative Adversarial Network*
GARCH	*Generalized Autoregressive Conditional Heteroskedasticity*
GBM	*Gradient Boosting Machine*
GCS	*Capital Goods and Services* (Setor de Bens de Capital e Serviços)
GRU	*Gated Recurrent Unit*
HQIC	*Hannan-Quinn Information Criterion*
ICA	*Independent Component Analysis*
InS	*In-Sample* (Dentro da Amostra)
ISRA	*In-Sample Recommendation Algorithm*
KNN	*K-Nearest Neighbors*
KWP	*Kelly Weighted Portfolio*
LASSO	*Least Absolute Shrinkage and Selection Operator*
LR	*Logistic Regression*
LSTM	*Long Short-Term Memory*

MACD	*Moving Average Convergence/Divergence*
MAE	*Mean Absolute Error*
MAO	*Moving Average Oscillator*
MAP	*Mean Average Precision*
MAPE	*Mean Absolute Percentage Error*
MCDM	*Multi-Criteria Decision Making* (Tomada de Decisão Multicritério)
MDD	*Maximum Drawdown*
MEMD	*Multivariate Empirical Mode Decomposition*
MFI	*Money Flow Index*
MGARCH	*Multivariate Generalized Autoregressive Conditional Heteroskedasticity*
MI	*Mass Index*
ML	*Machine Learning*
MLE	*Maximum Likelihood Estimation*
MLP	*Multilayer Perceptron*
MNB	*Multinomial Naive Bayes*
MOM	*Momentum*
MRR	*Mean Reciprocal Rank*
mRMR	*Minimum Redundancy Maximum Relevance*
MSR	*Modified Sharpe Ratio*
NB	*Naive Bayes*
NBIAS	*Normalized Bias*
NCO	*Net Change Oscillator*
NDCG	*Normalized Discounted Cumulative Gain*
OBV	*On Balance Volume*
OGB	*Oil, Gas and Biofuels* (Setor de Óleo, Gás e Biocombustíveis)
OHLC	*Open, High, Low, Close* (preços de abertura, máximo, mínimo e fechamento)

OHLCV	*Open, High, Low, Close, Volume*
OutS	*Out-of-Sample* (Fora da Amostra)
PCA	*Principal Component Analysis*
PO	*Price Oscillator*
PR	*Precision Rate*
PRBF	*Pearson Redundancy Based Filter*
PSY	*Psychology Line*
PT	*Prospect Theory*
RBM	*Restricted Boltzmann Machine*
RBF	*Radial Basis Function*
REM	*Robust Expectation Maximization*
RF	*Random Forest*
RMSE	*Root Mean Square Error*
RMI	*Relative Momentum Index*
RNA	Redes Neurais Artificiais
RNN	*Recurrent Neural Network*
ROC	*Rate of Change*
RPWP	*Risk Parity Weighted Portfolio*
RR	*Recall Rate*
RSI	*Relative Strength Index*
RSV	*Raw Stochastic Value*
SA	*Simulated Annealing*
SAE	*Stacked Autoencoder*
SARIMA	*Seasonal Autoregressive Integrated Moving Average*
SMA	*Simple Moving Average*
SMOTE	*Synthetic Minority Oversampling Technique*

SONAR	*SONAR Oscillator*
SONSIG	*SONAR Signal Line*
SOR	*Sortino Ratio*
SR	*Sharpe Ratio*
SROC	*Smoothed Rate of Change*
SSO	*Social Spider Optimization*
SVM	*Support Vector Machine*
SVMD	*Successive Variational Mode Decomposition*
SVR	*Support Vector Regression*
TEMA	*Triple Exponential Moving Average*
TesD	Conjunto de dados de teste (*Testing Dataset*)
TFT	*Temporal Fusion Transformer*
TGARCH	*Threshold Generalized Autoregressive Conditional Heteroskedasticity*
TODIM	Tomada de Decisão Interativa Multicritério
TraD	Conjunto de dados de treinamento (*Training Dataset*)
TRIX	*Triple Exponential Moving Average Oscillator*
TS	*Trading Signals* (Sinais de Negociação)
TU	Teoria da Utilidade
UTI	*Utilities* (Setor de Serviços Públicos)
VAR	*Vector Autoregression*
VIF	*Variance Inflation Factor*
VMA	*Volume Moving Average*
VMD	*Variational Mode Decomposition*
VOLA	*Chaikin Volatility*
VOS	*Volume Oscillator*
VROC	*Volume Rate of Change*

WFA	*Walk-Forward Analysis*
WR	*Win Rate*
WR%	*Williams %*
XGB	*Extreme Gradient Boosting*

Lista de Símbolos

EU	Utilidade esperada de uma alternativa
p_i	Probabilidade de ocorrência do resultado x_i
x_i	i-ésimo resultado possível de uma alternativa
$u(x_i)$	Utilidade associada ao resultado x_i
A_i	i-ésima alternativa (modelo de negociação candidato)
C_j	j-ésimo critério de avaliação (indicador de desempenho)
d_{ij}	Valor de desempenho do modelo i no indicador j
D	Matriz de decisão composta pelos elementos d_{ij}
M	Número de critérios (indicadores) na matriz de decisão
N	Número de alternativas (modelos candidatos) ou de ações
w	Vetor de pesos atribuídos aos critérios
w_r	Peso de referência (maior peso entre os critérios)
w_{cr}	Razão entre o peso do critério c e o peso de referência r
n_{ij}	Valor normalizado da matriz de decisão para o modelo i no critério j
θ	Fator de atenuação de perdas na função de dominância do TODIM
α, β	Coeficientes de concavidade e convexidade da função de valor (Teoria do Prospecto)
λ	Grau de aversão à perda na função de valor
$\Phi_j(A_i, A_k)$	Dominância parcial do modelo i sobre o modelo k no critério j
$\Phi(A_i, A_k)$	Dominância global do modelo i sobre o modelo k
$\Phi(A_i)$	Desempenho global do modelo de negociação i
$\xi(A_i)$	Medida de dominância global normalizada (escore final de ranqueamento)
q_j	Importância relativa (peso bruto) do critério C_j antes da normalização

w_j	Peso normalizado do critério C_j; equivalente a w_c na formulação do Capítulo 2
$\phi_c(A_i, A_j)$	Dominância parcial da alternativa A_i sobre A_j no critério c (notação do Capítulo 2)
$\delta(A_i, A_j)$	Dominância global da alternativa A_i sobre A_j; equivalente a $\Phi(A_i, A_k)$ (notação do Capítulo 2)
d_1	Comprimento da janela de treinamento no WFA (em dias)
d_2	Comprimento da janela deslizante no WFA (em dias)
L	Número total de observações de um ativo
Q	Número de rodadas do WFA
H	Número de observações descartadas para alinhamento do WFA
p	Número de *features* (indicadores técnicos) utilizadas como entrada
TRD_t	Taxa de retorno diário do ativo no instante t
$TRMD_t$	Taxa de retorno diário do modelo de negociação no instante t
TS_t	Sinal de negociação gerado pelo modelo no instante t
P_t	Valor do portfólio no instante t
I_t	Valor do investimento no tempo t (no calculo do ART)
n_t	Valor líquido da carteira no instante t
h	Período de detenção da carteira
m	Número de períodos individuais em um ano (fator de anualização)
σ_h	Desvio padrão da taxa de retorno no período h
r_f	Taxa de retorno livre de risco (*benchmark*)
D_τ	*Drawdown* calculado no tempo τ
r	Taxa de retorno do portfólio
τ	Taxa de retorno mínima aceitável (no cálculo do SOR)
S	Assimetria (*skewness*) da distribuição de retornos
K	Número total de modelos candidatos nas métricas de generalização

Ks	Curtose (*kurtosis*) da distribuição de retornos
T_i	i-ésimo período de recuperação de um *drawdown*
H_t	*High Water Mark* (pico histórico do valor do portfólio até t)
ρ	Coeficiente de correlação de Pearson entre dois atributos
C_t	Preço de fechamento (*Close*) no dia t
$H_t^{\max}$	Preço máximo nos últimos n dias
$L_t^{\min}$	Preço mínimo nos últimos n dias
O_t	Preço de abertura (*Open*) no dia t
V_t	Volume de negociação no dia t
n	Período (em dias) utilizado no cálculo dos indicadores técnicos
Z	número total de retornos na série (no cálculo do SOR)
R_n^k	Conjunto dos k melhores modelos recomendados no cenário de calibração para o n-ésimo setor
T_n^k	Conjunto dos k melhores modelos recomendados no cenário de validação para o n-ésimo setor
k	Utilizado como top-k modelos e atributos
w_i	Peso do i-ésimo *spread* na composição do portfólio
σ_i	Desvio padrão (volatilidade) do i-ésimo *spread*
$\bar{r}$	Retorno médio diário do *spread* utilizado no cálculo do Índice de Sharpe
T	Fator de anualização utilizado no cálculo do Índice de Sharpe dos *spreads* ($T = 250$)

Sumário

1 Introdução

A composição de carteiras no mercado acionário frequentemente envolve ativos de diferentes setores econômicos, que podem apresentar padrões de comportamento, fatores de risco e respostas distintas a eventos macroeconômicos. Ao mesmo tempo, o mercado de ações é caracterizado por elevada incerteza, baixa relação sinal-ruído e dinâmica potencialmente não linear, o que dificulta a construção de modelos de negociação com desempenho consistente em todos os ativos. Nesse cenário, modelos de aprendizado de máquina têm sido empregados para apoiar a geração de sinais de negociação e a identificação de padrões em séries financeiras.

Entretanto, estudos indicam que um mesmo modelo pode apresentar variações significativas de desempenho quando aplicado a ações de diferentes setores (LV et al., 2019). Assim, recomendar o modelo de negociação mais adequado para cada setor ou ação torna-se uma etapa relevante no desenvolvimento de estratégias de *trading* no mercado acionário. De acordo com Lv et al. (2023), embora diversos estudos tenham investigado a aplicação de modelos de aprendizado de máquina à previsão no mercado acionário, poucos trabalhos abordam explicitamente a seleção sistemática do modelo mais adequado para diferentes setores da economia, considerando simultaneamente múltiplos indicadores de desempenho e de risco.

Apesar desses avanços, observa-se uma lacuna na literatura quanto à recomendação sistemática de modelos de *trading* para o mercado acionário brasileiro, especialmente quando se consideram simultaneamente a heterogeneidade setorial, múltiplos critérios de desempenho e de risco, e a capacidade de generalização fora da amostra. Além disso, ainda são escassos os estudos que comparam, de forma estruturada, recomendações em diferentes níveis de granularidade, como a recomendação por setor econômico e a recomendação individual por ação. Dessa forma, permanece em aberto a questão de como selecionar, entre diferentes modelos de aprendizado de máquina, aquele mais adequado para cada setor ou ativo da B3 (Brasil, Bolsa Balcão), considerando retorno, risco e robustez em dados fora da amostra.

1.1 Motivação e Relevância

A motivação deste estudo decorre da necessidade de apoiar a seleção de modelos de negociação aplicáveis a ações de diferentes setores da economia. No mercado acionário, ativos de um mesmo setor tendem a compartilhar fatores de risco, padrões de comportamento e reações semelhantes a eventos econômicos, políticos e conjunturais. Além disso, fenômenos comportamentais, como o chamado “efeito manada”, podem contribuir para

movimentos de rotação setorial e para o agrupamento de ativos com características semelhantes. Dessa forma, considerar as particularidades setoriais torna-se relevante para a recomendação de modelos de *trading*.

A seleção desses modelos também envolve desafios na avaliação de desempenho. O mesmo algoritmo pode apresentar resultados distintos conforme o ativo ou setor analisado, e diferentes métricas de retorno e de risco podem produzir classificações conflitantes. Por exemplo, um modelo pode apresentar maior taxa de acerto (do inglês *win rate*) e maior índice de Sharpe anualizado (do inglês *annualized Sharpe ratio*), enquanto outro pode apresentar maior taxa de retorno anualizado (do inglês *annualized return rate*). Nesse contexto, a definição do modelo mais adequado não deve depender de um único indicador, mas de uma abordagem capaz de ponderar múltiplos critérios simultaneamente.

Outro aspecto relevante é a capacidade de generalização dos modelos recomendados. Em aplicações financeiras, um modelo pode apresentar desempenho satisfatório nos dados de treinamento ou de validação, mas não manter esse desempenho nos dados fora da amostra. Assim, avaliar a robustez das recomendações em cenários fora da amostra (em inglês, *out-of-sample*) é fundamental para verificar sua aplicabilidade em estratégias de investimento.

Diante desse contexto, este trabalho propõe uma abordagem de recomendação de modelos de negociação para ações da B3, fundamentada na Tomada de Decisão Multicritério (em inglês, *Multi-Criteria Decision Making*, MCDM). Inicialmente, diferentes algoritmos de aprendizado de máquina são considerados candidatos a modelos para a geração de sinais de negociação. Em seguida, realizam-se testes retrospectivos (*backtesting*) com os modelos treinados, a fim de obter indicadores de desempenho e de risco das estratégias. Por fim, a partir do cálculo das vantagens relativas entre os modelos candidatos, aplica-se o método TODIM, acrônimo de Tomada de Decisão Interativa Multicritério, para recomendar o modelo mais adequado a cada setor ou ação analisada. O TODIM é um método MCDM consolidado e amplamente reconhecido (GOMES; LIMA, 1991; GOMES; LIMA, 1992), baseado nos fundamentos da teoria do prospecto e capaz de incorporar a aversão à perda no processo de decisão.

Assim, a relevância prática da pesquisa está associada ao apoio à elaboração de estratégias quantitativas de negociação e à composição de carteiras de ações de diferentes setores da economia. Do ponto de vista científico, a contribuição reside na modelagem da seleção de modelos de *trading* como um problema multicritério, integrando aprendizado de máquina, indicadores de retorno e risco, análise comportamental e avaliação de generalização fora da amostra.

1.2 Caracterização do Problema

A previsão de séries temporais no mercado financeiro e a identificação de padrões lucrativos constituem tarefas desafiadoras devido à natureza dos dados financeiros. O comportamento dos preços é influenciado por fatores econômicos, políticos, sociais e comportamentais, o que torna complexa a extração de padrões latentes e dificulta a generalização de algoritmos de aprendizado de máquina para diferentes ativos (LV et al., 2023).

Nesse contexto, o problema investigado nesta dissertação consiste em selecionar, entre diferentes modelos candidatos de aprendizado de máquina, aquele mais adequado para orientar estratégias de negociação em ações da B3. Essa seleção é dificultada pelo fato de que um mesmo modelo pode apresentar desempenho superior em um critério e inferior em outro. Assim, a definição do melhor modelo não pode depender de uma única métrica, o que exige uma metodologia capaz de combinar múltiplos indicadores de retorno e de risco (LV et al., 2023).

Trabalhos anteriores mostraram que a previsão de séries temporais pode ser uma ferramenta útil para especular sobre os retornos das ações (GONZALO; PITARAKIS, 2002). Técnicas de inteligência artificial podem ser aplicadas a tarefas de previsão de direção e/ou de valor de ativos financeiros, em razão de sua capacidade de identificar padrões difíceis de observar para analistas humanos (KORCZAK; HERNES, 2017). Ademais, essas técnicas tornam-se especialmente relevantes em cenários de elevada instabilidade, como períodos de crise, volatilidade pós-pandemia e tensões geopolíticas (WANG; WANG, 2021). Nessas situações, a eficiência dos mercados financeiros pode se deteriorar temporariamente, abrindo espaço para padrões mais previsíveis no curto prazo (IZZELDIN et al., 2023). Estudos indicam que, nesses contextos, os preços podem apresentar maior autocorrelação, persistência e rupturas na aleatoriedade, sinalizando desvios da hipótese dos mercados eficientes (SHAHID, 2022).

Adicionalmente, observa-se, em escala global, o crescente interesse e a maior participação de pessoas físicas no mercado financeiro. A globalização dos mercados, o aumento da renda das famílias em países em desenvolvimento (KANIEL; SAAR; TITMAN, 2008), a facilidade de acesso à informação e o avanço das tecnologias financeiras têm contribuído para ampliar a participação de investidores individuais nas bolsas de valores (United Nations Conference on Trade and Development, 2020). Um estudo da XP Investimentos (XP Investimentos, 2023) aponta que o número de pessoas físicas que investem em ações está diretamente relacionado à estabilidade do mercado e à qualidade de vida nos países analisados. Na bolsa brasileira, registrou-se um aumento de mais de 2,8 milhões de investidores individuais em 2020, o que representa um crescimento de 93,7% em relação a 2019 (B3 – Brasil, Bolsa, Balcão, 2023). A Figura 1 ilustra a porcentagem de pessoas que investiam em ações em diferentes países em 2020, destacando que, nos Estados Unidos,

aproximadamente 55% da população investia em ações, enquanto no Brasil esse percentual era de cerca de 3%.

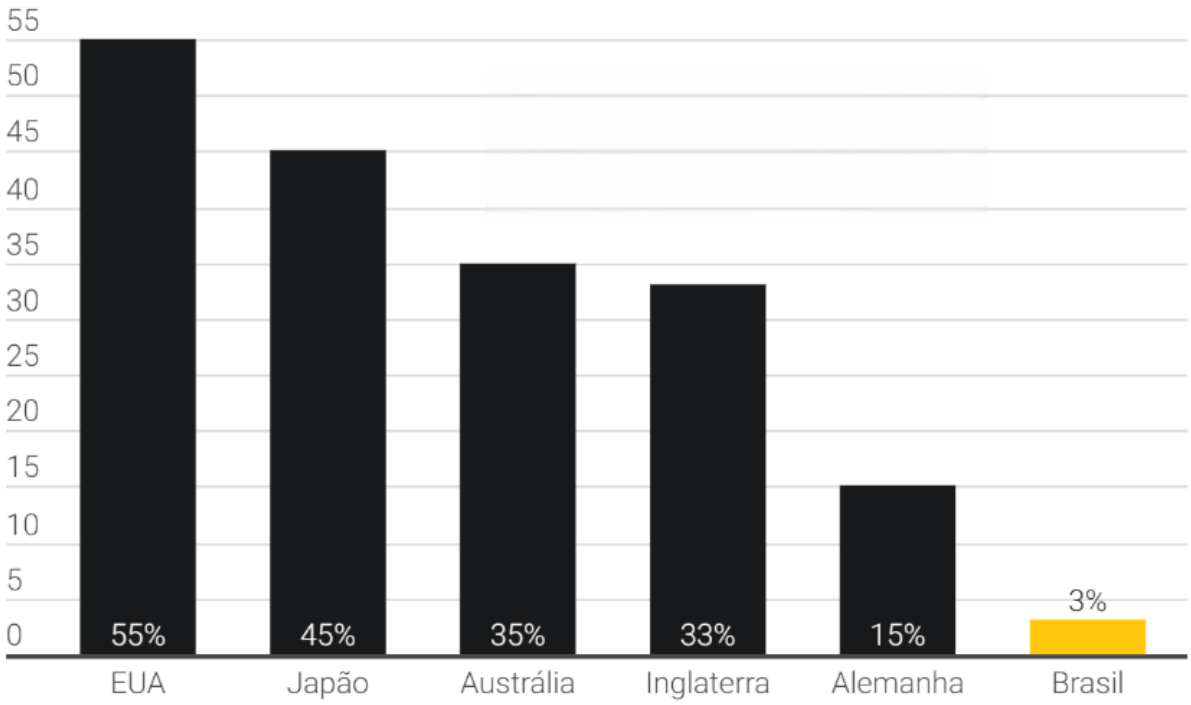

Figura 1 – Porcentagem de pessoas por país que investem em ações ano 2020.

Adicionalmente, um estudo recente, publicado em 2024, relatou a existência de 19,1 milhões de pessoas físicas distintas (CPFs únicos) que investiram na B3 em dezembro de 2023 (XP Investimentos, 2024). Considerando que a maior parte desses investidores é residente no Brasil e que, em 2023, a população brasileira era de aproximadamente 203 milhões (Instituto Brasileiro de Geografia e Estatística, 2023), esse contingente correspondia a cerca de 9,4% da população naquele ano.

Diante desse cenário, este trabalho propõe modelar a seleção de estratégias de investimento como um processo de tomada de decisão multicritério. Especificamente, sugere-se um algoritmo de recomendação de modelos de negociação de ações setoriais e individuais baseado no método TODIM. A abordagem busca integrar múltiplos indicadores de desempenho e de risco, permitindo identificar o algoritmo de aprendizado de máquina mais adequado para cada setor e ação, em vez de adotar uma solução única para todo o mercado.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral desta pesquisa é desenvolver e avaliar uma abordagem multicritério, baseada no método TODIM, para recomendar modelos de aprendizado de máquina aplicados ao *trading* de ações no mercado brasileiro.

1.3.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- realizar uma revisão da literatura sobre previsão de séries financeiras e modelos de aprendizado de máquina aplicados ao mercado de ações;
- investigar a hipótese de que não existe um modelo único de *trading* adequado a todos os setores econômicos, avaliando diferenças de desempenho entre setores;
- avaliar o impacto da seleção de atributos, por meio de correlação de Pearson e informação mútua (*mutual information*), no desempenho dos modelos de *trading*;
- aplicar o método TODIM à recomendação de modelos de *trading*, incorporando múltiplos critérios de retorno e risco ao processo de decisão;
- comparar duas granularidades de recomendação: setorial, com indicação de um modelo ótimo por setor, e individual, com indicação de um modelo ótimo por ação;
- avaliar o desempenho de estratégias de portfólio baseadas nos modelos recomendados, comparando-as à estratégia Segurar e Manter (do inglês *Buy-and-Hold*) e ao índice Ibovespa, em termos de retorno e risco.

1.4 Contribuições da Dissertação

As principais contribuições desta dissertação são:

- adaptação da estrutura de recomendação de modelos de *trading* proposta por Lv et al. (2023) ao mercado acionário brasileiro (B3), contemplando a geração de sinais de negociação, a realização de *backtesting* e a recomendação multicritério baseada no método TODIM;
- avaliação do impacto da seleção de atributos no desempenho dos modelos de *trading*, utilizando um processo baseado em correlação de Pearson e informação mútua;
- proposição de uma abordagem de recomendação individual por ação, em complemento à recomendação setorial, permitindo comparar diferentes níveis de granularidade na seleção de modelos;
- adaptação de métricas de sistemas de recomendação e recuperação de informação para avaliar a capacidade de generalização das recomendações em cenários dentro e fora da amostra;

- desenvolvimento e avaliação de estratégias de portfólio baseadas nos modelos recomendados, incluindo comparações com a estratégia *Buy-and-Hold*, o índice Ibovespa e portfólios *Long-Short*;
- realização de uma avaliação experimental com ações da B3 distribuídas em diferentes setores econômicos, considerando múltiplos modelos candidatos de aprendizado de máquina.

1.5 Organização do Trabalho

O presente trabalho está organizado em seis capítulos. O Capítulo 1 apresenta a introdução, a motivação, a caracterização do problema, os objetivos e as principais contribuições da pesquisa. O Capítulo 2 aborda os fundamentos teóricos relacionados às séries temporais financeiras, à economia comportamental, à teoria do prospecto, aos indicadores técnicos e ao método TODIM. O Capítulo 3 apresenta a revisão da literatura, discutindo trabalhos relacionados à previsão no mercado de ações, à aplicação de aprendizado de máquina em finanças e à seleção de alternativas candidatas por meio de métodos multicritério. O Capítulo 4 descreve a abordagem proposta, incluindo a geração de sinais de negociação, o processo de *backtesting* e a recomendação multicritério de modelos de *trading*. O Capítulo 5 apresenta e discute os resultados experimentais obtidos com a aplicação da abordagem proposta ao mercado acionário brasileiro. Por fim, o Capítulo 6 reúne as considerações finais, as limitações do estudo e as sugestões para trabalhos futuros.

Cumpre registrar que, durante a elaboração desta dissertação, ferramentas de Inteligência Artificial generativa, especificamente o modelo *Claude Sonnet 5*, foram utilizadas como recurso de apoio à escrita. O seu uso restringiu-se estritamente à revisão ortográfica e gramatical, ao auxílio na tradução de termos específicos e à estruturação de tópicos para organização das ideias. O autor revisou cuidadosamente todo o conteúdo gerado ou alterado pela ferramenta e assume total e integral responsabilidade pela precisão, originalidade e integridade do texto final apresentado.

2 Fundamentação Teórica

Este capítulo apresenta os fundamentos teóricos que sustentam a metodologia proposta. Inicialmente, discutem-se conceitos de finanças comportamentais, com ênfase na Teoria da Utilidade Esperada e na Teoria do Prospecto, que fornece a base conceitual para o método TODIM. Em seguida, apresentam-se os métodos de decisão multicritério e a formulação do TODIM, destacando-se como a assimetria entre ganhos e perdas é incorporada ao processo de ranqueamento. Por fim, discutem-se os fundamentos da análise técnica e os indicadores técnicos utilizados como variáveis de entrada dos modelos de aprendizado de máquina, bem como os cuidados para evitar viés de antecipação e instabilidades numéricas.

2.1 Finanças Comportamentais

As teorias financeiras tradicionais, em especial as baseadas na racionalidade dos agentes, assumem que os indivíduos tomam decisões de forma consistente, avaliando as alternativas disponíveis e escolhendo aquela que maximiza sua utilidade esperada (PINDYCK; RUBINFELD, 2013). Sob essa perspectiva, o decisor é capaz de processar as informações relevantes, comparar riscos e retornos e selecionar a alternativa mais adequada aos seus objetivos. Entretanto, diversas evidências empíricas indicam que os indivíduos nem sempre se comportam de forma plenamente racional. Em situações de risco e incerteza, decisões financeiras podem ser influenciadas por vieses cognitivos, emoções, heurísticas, aversão à perda e pela forma como as alternativas são apresentadas. Nesse contexto, as finanças comportamentais surgem como uma abordagem complementar às teorias tradicionais, ao incorporar fatores psicológicos à análise do processo decisório. A Teoria do Prospecto, proposta por Kahneman e Tversky, constitui um dos principais pilares das finanças comportamentais. Essa teoria fornece uma estrutura descritiva para compreender como os indivíduos avaliam ganhos, perdas e probabilidades em contextos de risco, sendo especialmente relevante para este trabalho por fundamentar o método TODIM, posteriormente utilizado na recomendação de modelos de negociação (KAHNEMAN; TVERSKY, 1979).

2.1.1 Teoria da Utilidade Esperada

Nesta subseção apresenta-se uma breve revisão da literatura sobre a Teoria da Utilidade Esperada, sem a pretensão de esgotar a profundidade do tema. Na teoria econômica clássica, os preços são determinados pelos custos de produção. Já na teoria econômica neoclássica, outros agentes passam a influenciar os preços dos produtos, destacando-se os conceitos de oferta e demanda. A demanda resulta da escolha de um ou mais indivíduos

de adquirir determinado produto em detrimento de outro, introduzindo, assim, o conceito de escolha. Na Teoria da Utilidade, a esse conceito somam-se os princípios de racionalidade e de utilidade, entendidos como a satisfação que o indivíduo obtém em função do resultado de sua escolha (YOSHINAGA; RAMALHO, 2014).

A Teoria da Utilidade (TU) tem suas origens nos trabalhos de Pascal e Fermat, que, pela primeira vez, buscaram analisar o comportamento humano por meio de ferramentas matemáticas. Décadas mais tarde, Bernoulli propôs, ainda que de forma rudimentar, os fundamentos da Teoria da Utilidade ao estudar o Paradoxo de São Petersburgo. Neste trabalho, introduziu-se o conceito de utilidade, distinto do valor monetário, entendido como um valor de natureza "moral" ou associado ao nível de satisfação (SCHMIDT, 2004).

O conceito de utilidade, incorporado à TU, corresponde à satisfação que o tomador de decisão obtém das consequências de sua escolha. Dessa forma, a teoria incorpora as preferências individuais ao processo decisório, uma vez que a utilidade tende a variar de indivíduo a indivíduo. Além disso, essa utilidade tende a diminuir à medida que mais unidades de um bem são consumidas, fenômeno conhecido como utilidade marginal decrescente (PINDYCK; RUBINFELD, 2013).

A TU também introduz os conceitos de aversão ao risco e de comportamento de busca por risco, que influenciam diretamente a postura do indivíduo diante das escolhas. Ao mesmo tempo, a teoria pressupõe que o tomador de decisão é racional e fará escolhas coerentes com base em todas as informações disponíveis.

A Teoria da Utilidade Esperada constitui uma das principais bases da teoria econômica para a análise da tomada de decisão sob risco. De acordo com essa abordagem, o indivíduo racional escolhe, entre diferentes alternativas incertas, aquela que maximiza sua utilidade esperada. A utilidade representa o nível de satisfação associado a um determinado resultado e não é necessariamente equivalente ao valor monetário obtido.

Formalmente, considerando um conjunto de possíveis resultados x_i, cada um associado a uma probabilidade p_i, a utilidade esperada de uma alternativa pode ser representada por:

$$EU = \sum_{i=1}^{n} p_i u(x_i), \tag{1}$$

em que EU representa a utilidade esperada, p_i corresponde à probabilidade de ocorrência do resultado x_i, e $u(x_i)$ representa a utilidade associada a esse resultado.

Quando se trata do conceito de risco, este pode ser interpretado como uma loteria, em que o resultado é incerto. É amplamente aceito que a maioria das pessoas apresenta aversão ao risco. Um exemplo clássico ocorre quando um indivíduo deve escolher entre receber um valor com cem por cento de certeza e o mesmo valor com probabilidade

de não recebê-lo; em geral, a escolha recai sobre a alternativa sem risco (CRAINICH; EECKHOUDT; MENEGATTI, 2019).

Quando existem duas alternativas, sendo a primeira associada a um valor menor e sem risco, e a segunda a um valor maior, porém arriscado, a decisão do investidor dependerá da função de utilidade associada a cada resultado. A escolha entre essas opções reflete o grau de aversão ao risco do indivíduo. Na Teoria da Utilidade, essa aversão é definida como consequência da diminuição da utilidade marginal da riqueza monetária (CRAINICH; EECKHOUDT; MENEGATTI, 2019).

Apesar de sua ampla utilização para modelar escolhas sob risco, a Teoria da Utilidade Esperada apresenta limitações quando confrontada com o comportamento observado em situações reais. Na prática, os indivíduos nem sempre dispõem de todas as informações relevantes, podem interpretá-las de forma incompleta e frequentemente tomam decisões sob restrições cognitivas, emocionais e temporais. Assim, a hipótese de racionalidade plena torna-se limitada para explicar escolhas efetivas em ambientes complexos e incertos (ZHANG; LI; GUO, 2018).

Outra crítica relevante refere-se à suposição de racionalidade do indivíduo no processo decisório. A teoria pressupõe que o tomador de decisão segue um processo coerente e consistente. No entanto, na prática, os indivíduos tendem a ser contraditórios, influenciados por vieses e sujeitos a mudanças frequentes de comportamento, o que resulta, por exemplo, em oscilações no valor de utilidade atribuído a um ativo (ZHANG; LI; GUO, 2018).

Outras discrepâncias entre a teoria e a prática nos mercados financeiros dizem respeito às dinâmicas de seleção e ao tratamento do tempo. Na teoria, essas decisões são frequentemente modeladas de forma estática, enquanto, na realidade, ocorrem ao longo de diferentes horizontes temporais. Ademais, os retornos projetados de ativos arriscados são, em muitos modelos tradicionais, assumidos como normalmente distribuídos. Na prática, entretanto, séries financeiras frequentemente apresentam assimetria, caudas pesadas, volatilidade temporal variável e influência de choques externos, o que limita a adequação da hipótese de normalidade (ZHANG; LI; GUO, 2018).

Essas limitações motivaram o desenvolvimento de modelos alternativos mais próximos da realidade observada. Como resultado, emergiram teorias que preservam alguns fundamentos da Teoria da Utilidade, mas incorporam fatores psicológicos ao processo decisório, reconhecendo o caráter humano, e não estritamente racional, dos tomadores de decisão.

2.1.2 Teoria do Prospecto

É importante compreender a Teoria do Prospecto (do inglês *Prospect Theory*, PT), uma vez que o método TODIM a incorpora em sua metodologia. De acordo com Kahneman

e Tversky (1979), a Teoria do Prospecto apresenta-se como um modelo descritivo das escolhas sob risco, proposto como substituto e como crítica à Teoria da Utilidade Esperada na explicação dos comportamentos observados.

Como resultado dessas críticas, diversos estudos passaram a evidenciar anomalias e inconsistências nas premissas básicas de racionalidade plena adotadas pelos modelos tradicionais de decisão. Nesse contexto, uma nova teoria ganhou destaque entre alguns pesquisadores ao incorporar não apenas o aspecto da racionalidade, mas também elementos psicológicos do tomador de decisão, que, por vezes, se mostra imprevisível, conflitivo e inconstante (YOSHINAGA; RAMALHO, 2014).

Segundo Kahneman e Tversky (1979), alguns argumentos centrais justificaram a criação da PT, ao demonstrar que a Teoria da Utilidade falhou em modelar adequadamente o comportamento de escolha dos indivíduos em diferentes situações e que, em contextos específicos, diversos de seus axiomas foram violados. A Teoria da Utilidade busca prescrever como as pessoas deveriam agir, mas não consegue explicar como elas efetivamente se sentem nem quais valores atribuem às alternativas disponíveis. Por se basear exclusivamente na racionalidade, essa teoria tem dificuldade em abarcar a irracionalidade inerente às emoções, o que acaba por influenciar o processo decisório.

Ao proporem a Teoria do Prospecto, Kahneman e Tversky, psicólogos de formação, buscaram contornar as limitações associadas ao conceito do homem estritamente racional (KAHNEMAN; TVERSKY, 1979). Por meio de uma série de experimentos, demonstrou-se que, em muitos casos, o investidor ou tomador de decisão se comporta como um ser humano dotado de preferências e preconceitos intrínsecos, sendo influenciado por esses fatores mesmo diante de evidências contrárias. Ademais, os indivíduos não são plenamente capazes de processar e absorver toda a informação disponível, nem de compreender integralmente as correlações entre os fatores no momento da tomada de decisão (CURTIS, 2004).

Os resultados apresentados em Kahneman e Tversky (1979) mantêm alinhamento com os fundamentos da Teoria da Utilidade Esperada, mas revelam diferenças significativas quando se analisam situações sob a perspectiva de perdas e ganhos em ambientes de risco. No domínio das perdas, o indivíduo tende a comportar-se como um buscador de risco, preferindo uma aposta a uma perda certa; no domínio dos ganhos, por sua vez, o mesmo indivíduo tende a apresentar comportamento avesso ao risco, optando por um ganho menor e garantido em detrimento de um ganho maior associado à incerteza.

Os autores demonstraram, ainda, que o indivíduo é influenciado não apenas pelo seu julgamento lógico, mas também por fatores comportamentais presentes no momento da decisão. Em situações de loteria, as pessoas mostraram-se mais propensas a assumir riscos quando confrontadas com a possibilidade de perdas maiores do que as de uma perda certa. Em contrapartida, quando a escolha envolvia ganhos maiores com determinada

probabilidade, em comparação a um ganho menor, porém garantido, a tendência observada foi optar pela segunda alternativa. Esse comportamento contraria as previsões da Teoria da Utilidade, evidenciando que o mesmo tomador de decisão pode apresentar atitudes conflitantes quando confrontado com situações de ganho e de perda (SILVA et al., 2014).

A Teoria do Prospecto fundamenta-se em dois pontos-chave. O primeiro refere-se à forma da função de valor, que apresenta concavidade no domínio dos ganhos e convexidade no domínio das perdas, conforme ilustrado na Figura 2. O segundo ponto está relacionado à ponderação subjetiva das probabilidades, considerando que os indivíduos tendem a atribuir pesos decisórios distintos das probabilidades objetivas. Em geral, pequenas probabilidades podem ser superestimadas, enquanto probabilidades médias ou elevadas podem ser subestimadas, fenômeno que contribui para vieses sistemáticos no julgamento (TVERSKY; KAHNEMAN, 2016) *apud* (PUPPO et al., 2022).

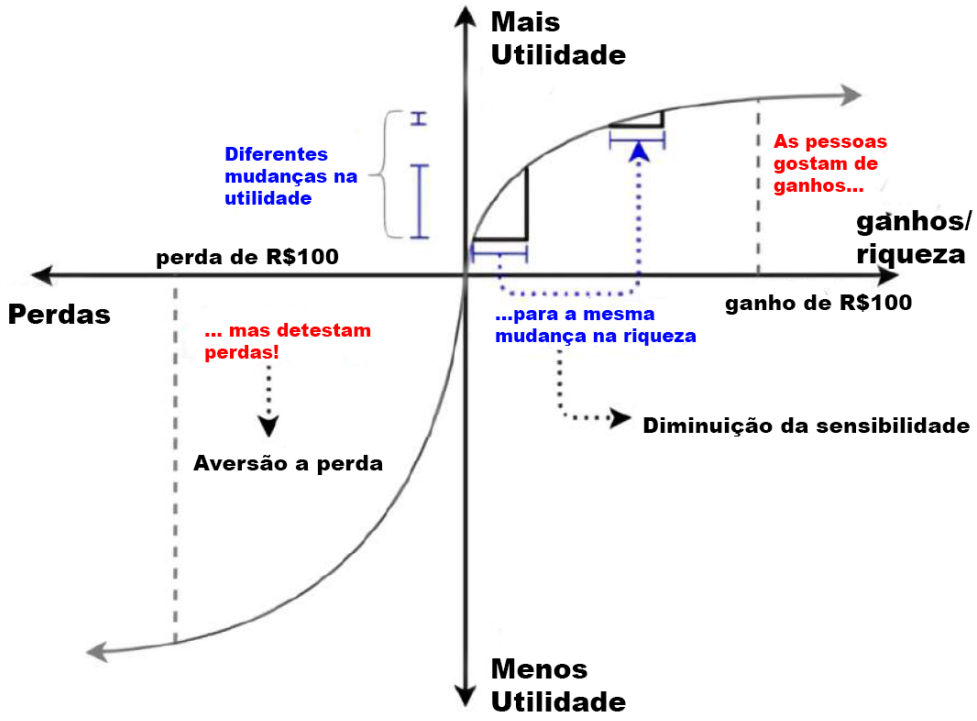

Figura 2 – Curva em S da Teoria do Prospecto. Sensibilidade decrescente: à medida que a riqueza cresce, o impacto de um determinado incremento de riqueza diminui. Aversão à perda: os indivíduos respondem de forma assimétrica a ganhos e perdas.

Adaptado de: Puppo et al. (2022).

A Figura 2 evidencia que os indivíduos apresentam percepções distintas quanto a ganhos e perdas. O formato diferenciado das curvas indica que a insatisfação associada a uma perda é maior do que a satisfação decorrente de um ganho de igual magnitude. Em outras palavras, perder uma quantia provoca um impacto emocional mais intenso do que o prazer sentido ao ganhá-la.

No mesmo estudo, Kahneman e Tversky (1979) identificaram outros aspectos relevantes que reforçam a influência do comportamento humano na tomada de decisão. Um deles mostra que o mesmo indivíduo pode tomar decisões distintas, e até opostas, diante do mesmo problema e com a mesma informação, desde que essa informação seja apresentada de forma diferente. Esse fenômeno é denominado efeito de enquadramento, ou *framing effect*, e demonstra que a forma de apresentação de uma alternativa pode alterar a percepção de risco e influenciar a decisão do indivíduo.

Um exemplo simples para ilustrar esse efeito é o de um paciente hospitalizado que precisa submeter-se a uma cirurgia. Caso o médico informe que existe 20% de chance de morte durante o procedimento, o paciente tende a concentrar sua atenção nessa possibilidade negativa, atribuindo-lhe um peso maior do que o racionalmente esperado, em função da forma como a informação foi apresentada. Por outro lado, se o mesmo médico informar que a cirurgia possui 80% de chance de sucesso, a tendência é que o paciente se concentre na elevada probabilidade de um desfecho positivo.

Esse exemplo demonstra que, embora se trate da mesma informação, a referência adotada altera significativamente a percepção do indivíduo sobre o valor e o risco envolvidos. Esse efeito também é discutido por Kahneman e Tversky (1979), que mostra que o mesmo problema, quando apresentado de formas distintas, pode levar a conclusões e expectativas diferentes, apesar de as probabilidades subjacentes serem idênticas.

Outro aspecto que diferencia a Teoria da Utilidade (TU) da Teoria do Prospecto (PT) diz respeito à função de valor. Enquanto a TU avalia o valor apenas em termos de utilidade associada ao resultado final, a PT incorpora explicitamente um ponto de referência na avaliação. Considere, por exemplo, dois indivíduos, A e B, com patrimônios iniciais de R$10.000,00 e R$60.000,00, respectivamente, que passam a possuir R$30.000,00 em um período posterior. Pela Teoria da Utilidade, ambos deveriam apresentar o mesmo nível de satisfação, uma vez que o valor final é o mesmo. No entanto, sem considerar o ponto de referência, não é possível captar que A obteve um ganho, enquanto B sofreu uma perda, o que implica percepções subjetivas distintas, apesar de ambos possuírem o mesmo valor final (PUPPO, 2024).

O ponto de referência é, portanto, fundamental, pois influencia tanto o comportamento quanto a percepção de valor dos indivíduos. Outro elemento relevante diz respeito à função de ponderação das probabilidades, que reflete a tendência humana de distorcer as probabilidades: valores muito baixos tendem a ser percebidos como nulos, enquanto valores muito elevados tendem a ser tratados como certos. Além disso, as probabilidades intermediárias podem ser superestimadas devido à dificuldade dos indivíduos em assimilar valores de forma linear (PUPPO, 2024).

Esse aspecto é amplamente explorado no setor de seguros. Por exemplo, embora a probabilidade de ocorrência de um acidente aéreo seja estatisticamente muito baixa,

indivíduos com medo de voar podem percebê-la como significativamente maior do que o risco objetivo. Nesses casos, a percepção subjetiva do risco pode tornar viável a contratação de seguros, uma vez que o risco percebido é superior ao risco estatístico (PUPPO, 2024).

Cabe destacar que, em 1992, Tversky e Kahneman (1992) apresentaram uma versão revisada da Teoria do Prospecto, conhecida como Teoria do Prospecto Cumulativa (do inglês *Cumulative Prospect Theory*, CPT). Essa formulação introduz uma abordagem baseada na ordenação para a avaliação de prospectos, preservando os conceitos de ponto de referência, função-valor em forma de S e aversão à perda, além de ampliar a teoria original para lidar com prospectos que envolvem múltiplos desfechos.

Por fim, ainda há debates na literatura sobre qual dos dois modelos, Teoria do Prospecto ou Teoria da Utilidade, é mais adequado para explicar o comportamento dos tomadores de decisão. Existem defensores de ambas as abordagens, bem como estudos que sugerem que a integração de elementos de ambas as teorias pode representar uma alternativa mais apropriada para modelar decisões em contextos reais (CURTIS, 2004).

Os conceitos discutidos nesta seção fornecem a base teórica para a aplicação do método TODIM, cuja formulação incorpora a assimetria entre ganhos e perdas no processo de avaliação das alternativas.

2.2 Métodos de Decisão Multicritério

2.2.1 Fundamentos dos Métodos Multicritério

Nesta subseção, discutem-se os métodos de decisão multicritério (em inglês, *Multi-Criteria Decision Making*, MCDM). Tal abordagem é necessária, pois o método TODIM é uma ferramenta de decisão multicritério, sendo fundamental compreender seu funcionamento, bem como suas vantagens e limitações.

Grande parte das atividades do cotidiano decorre de algum processo decisório. Em maior ou menor grau de complexidade, as decisões fazem parte da vida diária e ocorrem com recorrência. Entretanto, à medida que essas decisões se tornam mais complexas, surge a necessidade de empregar processos mais estruturados, capazes de organizar as informações e auxiliar na análise das alternativas disponíveis.

Quando essas decisões envolvem questões de maior escala, como a seleção de projetos, a avaliação de fornecedores ou a escolha de candidatos, passam a existir múltiplos objetivos, critérios e níveis de risco. Nesses casos, tornam-se necessárias ferramentas específicas que considerem, de forma sistemática, todas as informações relevantes no processo decisório.

Originários da área de Pesquisa Operacional, os métodos de decisão multicritério

foram desenvolvidos para auxiliar a tomada de decisão em ambientes de incerteza e risco. Em situações em que há *trade-offs*, isto é, quando uma alternativa ótima em um critério se torna inviável devido a conflitos com outros fatores, os métodos MCDM permitem que o tomador de decisão compreenda os impactos associados às escolhas realizadas e reordene suas prioridades de forma estruturada (VELASQUEZ; HESTER, 2013).

Os métodos de decisão multicritério permitem incorporar critérios conflitantes e avaliar e ponderar diferentes alternativas de acordo com Soare e Chinelli (2024). Essas ferramentas são amplamente utilizadas em problemas de ranqueamento e de avaliação sob critérios conflitantes (SIRIGIRI; HOTA; SHARMA, 2015), auxiliando na compreensão dos riscos envolvidos no processo decisório. Em geral, os métodos MCDM trabalham com valores incertos e incorporam aspectos relacionados à utilidade ou à função de valor, de acordo com as preferências do tomador de decisão, resultando em medidas de utilidade esperada para cada alternativa (WALLENIUS et al., 2008).

Os métodos MCDM permitem a análise de problemas altamente complexos, avaliando alternativas em múltiplas dimensões e identificando a solução mais adequada ao problema considerado (DOLAN, 2010). Esses modelos podem lidar tanto com atributos quantitativos quanto com qualitativos, o que os torna ferramentas robustas para situações complexas. Assim, além de avaliações numéricas, os métodos multicritério são capazes de incorporar julgamentos subjetivos ao processo de análise (TAHERDOOST; MADANCHIAN, 2023).

Segundo Zarghami e Szidarovszky (2011), os métodos MCDM consideram um conjunto finito de alternativas, que devem ser avaliadas com base em um conjunto limitado de critérios, como os físicos, econômicos, ambientais e sociais. Esses critérios são definidos de acordo com os objetivos e as preferências do tomador de decisão. Como resultado, o modelo fornece uma ordenação das alternativas, acompanhada de um valor de desempenho global associado a cada uma delas.

Os métodos de decisão multicritério são, em sua maioria, compensatórios, o que significa que um desempenho superior em determinado critério pode compensar um desempenho inferior em outro, equilibrando o valor global da alternativa. Contudo, esse tipo de abordagem nem sempre é apropriado. Em determinadas situações, torna-se necessário empregar modelos não compensatórios ou híbridos, nos quais a compensação entre critérios é limitada (ARMAN; ARMAN; HADI-VENCHEH, 2022).

Apesar de sua capacidade de ponderar múltiplas informações e produzir resultados mesmo diante de critérios conflitantes, os métodos MCDM dependem, invariavelmente, da intervenção do tomador de decisão. Essa intervenção ocorre, por exemplo, na definição dos pesos dos critérios, na seleção das alternativas ou na aceitação do resultado final. Dessa forma, os métodos permanecem sujeitos à subjetividade e às preferências individuais do tomador de decisão (ALJINOVIć; MARASOVIć; SESTANOVIć, 2021).

De acordo com Dolan (2010), os métodos de decisão multicritério podem ser entendidos como ferramentas que orientam o tomador de decisão na avaliação do desempenho de possíveis alternativas, considerando suas vantagens e desvantagens em cada critério analisado. Em problemas de elevada complexidade, é comum que o processo decisório seja conduzido por meio da decomposição do problema em subproblemas menores, que são avaliados individualmente e, posteriormente, recompostos em uma análise global.

Devido à grande variedade de modelos disponíveis, um dos principais desafios enfrentados pelo analista é identificar qual método é mais adequado ao problema em questão (TAHERDOOST; MADANCHIAN, 2023). Não é possível afirmar que exista um método ideal universal, uma vez que sua adequação depende fortemente das características do problema. Além disso, alguns modelos apresentam limitações que restringem sua aplicação a determinados contextos, o que tem impulsionado o desenvolvimento de modelos híbridos voltados a resolver problemas específicos que não são plenamente atendidos pelos métodos tradicionais (LEONETI; GOMES, 2021).

A escolha de um método de decisão multicritério é limitada, inicialmente, por sua capacidade de processar o número de alternativas envolvidas. Ademais, é fundamental que o analista compreenda a metodologia de cada modelo, bem como suas vantagens e limitações, para evitar que o processo decisório se torne uma “caixa-preta”. Essa compreensão permite ao usuário interpretar adequadamente os resultados obtidos (VELASQUEZ; HESTER, 2013).

Os problemas de seleção abordados pelos métodos MCDM são, em geral, complexos, pois envolvem múltiplos critérios qualitativos e quantitativos que frequentemente entram em conflito. O desafio consiste em encontrar uma solução que atenda às expectativas dos tomadores de decisão, cujas percepções podem ser enviesadas e não totalmente capturadas por métodos clássicos. Nesse contexto, ferramentas computacionais podem auxiliar o processo, permitindo o cálculo de parâmetros de difícil quantificação e a avaliação da robustez dos modelos aplicados, como a simulação de diferentes pesos dos critérios (KOU; PENG; WANG, 2014).

Adicionalmente, os métodos MCDM têm sido amplamente aplicados em diversas áreas de pesquisa. No contexto de alocação de portfólio, por exemplo, Alali e Tolga (2019) aplicaram o método TODIM e demonstraram que os portfólios obtidos apresentaram desempenho significativamente superior ao dos portfólios com pesos iguais. De forma semelhante, Lv et al. (2023) utilizaram o TODIM como algoritmo de recomendação para selecionar o modelo de *trading* mais adequado para diferentes setores industriais do mercado acionário. Já Kou et al. (2020) empregaram métodos de MCDM para avaliar técnicas de seleção de características em problemas de classificação de texto com conjuntos de dados reduzidos.

No campo das ciências sociais, Yu et al. (2021) propuseram uma abordagem baseada no método de maximização do desvio e no modelo TODIM para avaliar a vulnerabilidade

social na província de Jiangsu. Em outra aplicação, Wang et al. (2022) apresentaram uma abordagem integrada baseada em simulações, combinando o método *best-worst* e seu modelo não linear para determinar os pesos ideais dos critérios e, posteriormente, aplicando o método TODIM para classificar o desempenho de empresas petrolíferas nacionais.

Por fim, a crescente complexidade e incerteza dos problemas financeiros, intensificadas por fatores como a globalização, as crises econômicas e os avanços tecnológicos, têm tornado insuficientes os métodos tradicionais de análise, como os modelos econométricos e estatísticos. Diante de problemas que envolvem múltiplas variáveis de difícil mensuração, a aplicação de métodos de decisão multicritério tem se expandido, frequentemente em associação com técnicas de inteligência computacional (MARQUéS; GARCIA; SáNCHEZ, 2020).

No contexto desta dissertação, os métodos de decisão multicritério são essenciais, pois a escolha do modelo de negociação não pode basear-se em um único indicador, o que exige a consideração simultânea de múltiplos critérios de desempenho e de risco.

2.2.2 Tomada de Decisão Interativa Multicritério - TODIM

A escolha do método TODIM neste trabalho justifica-se por sua capacidade de incorporar a assimetria entre ganhos e perdas no processo decisório, característica particularmente relevante em problemas financeiros, nos quais o retorno e o risco são percebidos de forma não linear pelos investidores.

Considerando os avanços proporcionados pela Teoria do Prospecto na compreensão do processo decisório sob risco, desenvolveu-se o método TODIM com o objetivo de modelar a forma como o decisor manifesta suas preferências diante de situações de risco (GOMES; LIMA, 1992).

O principal diferencial do método TODIM reside em sua capacidade de integrar, em sua metodologia, o comportamento de racionalidade limitada do tomador de decisão, convertendo a sensibilidade a perdas e ganhos em uma função de valor (YIN et al., 2019).

O método utiliza funções distintas nas comparações pareadas: quando uma alternativa apresenta desempenho superior em determinado critério (ganho), aplica-se uma função específica; quando apresenta desempenho inferior (perda), aplica-se outra. Dessa forma, perdas e ganhos são ponderados de forma assimétrica (LV et al., 2023).

Diferentemente da maioria dos métodos multicritério, que assumem que o decisor sempre seleciona a alternativa com maior valor agregado segundo a Teoria da Utilidade, o TODIM adota uma abordagem inspirada na Teoria do Prospecto, incorporando a diferença relativa entre ganhos e perdas à medida que o valor global cresce (LEONETI; GOMES, 2021).

Por essas razões, o método despertou o interesse da comunidade científica, sendo um dos primeiros métodos MCDM a incorporar conceitos de comportamento individual, derivados da Teoria do Prospecto, em sua modelagem, além de apresentar capacidade de lidar com problemas complexos que envolvem múltiplas alternativas e critérios (WANG; LIU, 2017).

No contexto desta dissertação, a Teoria do Prospecto é particularmente relevante por modelar a avaliação assimétrica de ganhos e perdas em relação a um ponto de referência. O método TODIM operacionaliza essa assimetria no âmbito da decisão multicritério ao comparar alternativas par a par em cada critério e aplicar funções distintas para ganhos e perdas. Em termos práticos, essa característica permite penalizar desempenhos inferiores de forma diferente de como se recompensam desempenhos superiores, o que é compatível com problemas financeiros em que medidas de risco (perdas) tendem a exercer influência desproporcional sobre a preferência do decisor.

Considere um conjunto de alternativas A_i, $i = \{1,2,\ldots,n\}$, e um conjunto de critérios C_c, $c = \{1,2,\ldots,m\}$. A contribuição da alternativa A_i ao critério C_c é representada por x_{ic} (GOMES; LIMA, 1991).

Cabe destacar que, embora o modelo permita trabalhar com critérios quantitativos e qualitativos, é necessário converter todos os valores em valores numéricos. Um dos procedimentos mais utilizados para essa transformação é a aplicação da escala de Saaty (SAATY, 1990).

Como resultado, obtém-se a matriz de decisão $X = (x_{ic})_{n\times m}$. Em seguida, essa matriz é normalizada, resultando na matriz $P = (p_{ic})_{n\times m}$, em que p_{ic} representa o desempenho normalizado da alternativa A_i em relação ao critério C_c. A normalização é necessária para tornar comparáveis critérios expressos em diferentes escalas de medida.

O decisor ou especialista deve atribuir um vetor de pesos aos critérios, $w = [w_1,w_2,\ldots,w_m]$, devidamente normalizado, tal que:

$$\sum_{c=1}^{m} w_c = 1. \tag{2}$$

Define-se então um critério de referência C_r, cujo peso é dado por:

$$w_r = \max\{w_c \mid c = 1,2,\ldots,m\}. \tag{3}$$

Em geral, seleciona-se como critério de referência aquele que apresenta o maior peso (PUPPO, 2024).

Define-se w_{cr} como a razão entre o peso do critério c e o peso do critério de referência r, conforme:

$$w_{cr} = \frac{w_c}{w_r}. \tag{4}$$

Essa normalização permite expressar os pesos dos critérios em relação ao critério de referência, possibilitando a comparação dos desempenhos associados a diferentes critérios em uma mesma escala relativa.

Antes da aplicação da função de dominância, os critérios devem ser orientados na mesma direção de preferência. Assim, para critérios de benefício, valores maiores indicam melhor desempenho; para critérios de custo ou risco, a normalização deve ser realizada de modo que valores normalizados maiores também representem alternativas mais desejáveis. Esse procedimento evita interpretações equivocadas nas comparações pareadas entre alternativas.

A comparação entre as alternativas é realizada em pares, para cada critério. Considerando o desempenho da alternativa i no critério c como p_{ic} e o da alternativa j como p_{jc}, define-se:

$$\phi_c(A_i,A_j) = \begin{cases} \sqrt{\dfrac{w_{cr}(p_{ic}-p_{jc})}{\sum_{\ell=1}^{m} w_{\ell r}}}, & \text{se } p_{ic}-p_{jc} > 0, \\ 0, & \text{se } p_{ic}-p_{jc} = 0, \\ -\dfrac{1}{\theta}\sqrt{\dfrac{\left(\sum_{\ell=1}^{m} w_{\ell r}\right)(p_{jc}-p_{ic})}{w_{cr}}}, & \text{se } p_{ic}-p_{jc} < 0. \end{cases} \tag{5}$$

em que $\phi_c(A_i,A_j)$ representa o valor de dominância parcial da alternativa A_i sobre a alternativa A_j, associado ao critério c. O conjunto desses valores, calculado para todos os pares de alternativas, compõe a matriz de dominância parcial do critério analisado. Esse termo expressa a contribuição do critério c para a função de dominância global $\delta(A_i,A_j)$ (GOMES; LIMA, 1991).

Observações:

- $X = (x_{ic})_{n\times m}$ é a matriz de desempenho original das alternativas em relação aos critérios;
- $P = (p_{ic})_{n\times m}$ é a matriz de desempenho normalizada;
- $w = [w_1, \ldots, w_m]$ é o vetor de pesos dos critérios, e w_r é o maior peso, associado ao critério de referência;
- w_{cr} representa a razão entre o peso do critério c e o peso de referência r;
- θ é um parâmetro positivo que controla a penalização associada às perdas;
- $\delta(A_i,A_j)$ é a medida de dominância global da alternativa A_i sobre A_j, obtida pela agregação das dominâncias parciais.

A matriz de dominância global é definida por:

$$\delta(A_i,A_j) = \sum_{c=1}^{m} \phi_c(A_i,A_j), \qquad \forall(i,j). \tag{6}$$

Para obter o valor global de cada alternativa, essa matriz é normalizada, resultando em:

$$\xi_i = \frac{\sum_{j=1}^{n} \delta(A_i,A_j) - \min_k \sum_{j=1}^{n} \delta(A_k,A_j)}{\max_k \sum_{j=1}^{n} \delta(A_k,A_j) - \min_k \sum_{j=1}^{n} \delta(A_k,A_j)}, \qquad i \in \{1,2,\ldots,n\}. \tag{7}$$

Ou seja, para cada alternativa A_i, calcula-se a soma das dominâncias globais em relação a todas as demais alternativas do conjunto:

$$\sum_{j=1}^{n} \delta(A_i, A_j) = \delta(A_i, A_1) + \delta(A_i, A_2) + \cdots + \delta(A_i, A_n). \tag{8}$$

Esse valor representa o desempenho global acumulado da alternativa A_i: quanto maior, mais a alternativa i supera as demais no agregado de todos os confrontos pareados.

O índice k não é uma variável de soma; trata-se de um índice de busca. O somatório da Equação 7 é calculado para cada alternativa A_k do conjunto, e em seguida identificam-se o menor e o maior valor obtidos:

$$\min_k \sum_{j=1}^{n} \delta(A_k, A_j) = \text{menor somatório dentre todas as alternativas}, \tag{9}$$

$$\max_k \sum_{j=1}^{n} \delta(A_k, A_j) = \text{maior somatório dentre todas as alternativas}. \tag{10}$$

Para cada $k \in \{1, 2, \ldots, n\}$, computa-se o somatório correspondente, e o mínimo e o máximo são extraídos desse conjunto de n valores.

Portanto, a equação de ξ_i é uma normalização Min-Max aplicada ao somatório de A_i:

$$\xi_i = \frac{\text{(somatório de } A_i) - \text{(menor somatório de todas as alternativas)}}{\text{(maior somatório)} - \text{(menor somatório)}} \tag{11}$$

O resultado é $\xi_i \in [0,1]$ para toda alternativa i. A alternativa com o menor somatório recebe $\xi = 0$; a de maior somatório, $\xi = 1$; e todas as demais ficam proporcionalmente distribuídas nesse intervalo. A normalização garante que os escores finais sejam interpretáveis

e comparáveis independentemente da escala dos dados ou do número de alternativas avaliadas.

O valor ξ_i representa a medida de desejabilidade global, ou utilidade, da alternativa A_i. A ordenação das alternativas é obtida a partir dos valores de ξ_i, resultando em um ranking conforme as preferências do decisor ou do grupo de decisores. Alterações nos pesos dos critérios ou no parâmetro θ permitem realizar análises de sensibilidade (PUPPO et al., 2022).

A seguir, apresenta-se o pseudocódigo do algoritmo TODIM.

Algoritmo 1 Método TODIM

Entrada: Vetor de pesos w; conjunto de alternativas A; conjunto de critérios C; fator de atenuação das perdas θ
Saída: Ranking das alternativas

Inicialização
Construir matriz de decisão $\mathcal{X}$
Normalizar $\mathcal{X}$, obtendo a matriz $\mathcal{P}$
Definir o critério de referência C_r
Calcular os pesos relativos w_{cr}
for all (A_i, A_j), i,$j = 1, 2, \ldots, n$ **do**
    for all $c \in \{1, 2, \ldots, m\}$ **do**
        Calcular $\phi_c(A_i, A_j)$ conforme Eq. (5)
    end for
end for
Calcular $\delta(A_i, A_j)$ para todos os pares de alternativas, conforme Eq. (6)
Calcular ξ_i para todas as alternativas, conforme Eq. (7)
Ordenar as alternativas de acordo com ξ_i

2.3 Análise Técnica e Indicadores Técnicos

A análise técnica constitui uma abordagem metodológica voltada ao estudo do comportamento histórico dos preços de ativos financeiros, com o propósito de identificar padrões recorrentes que auxiliem a análise de possíveis movimentos futuros no mercado (ACHELIS, 2000). Suas raízes remontam à Teoria de Dow, formulada por Charles Dow por volta de 1900, da qual derivam princípios fundamentais como a natureza tendencial dos preços, a premissa de que os preços incorporam todas as informações disponíveis, bem como os conceitos de confirmação, divergência e suporte e resistência. Sob essa perspectiva, o preço de um ativo representa um consenso entre compradores e vendedores, sendo determinado, em grande medida, pelas expectativas dos participantes do mercado. Diferentemente da análise fundamentalista, que busca estimar o valor intrínseco de um ativo com base em projeções de lucros e variáveis macroeconômicas, a análise técnica

parte do pressuposto de que o comportamento passado dos preços fornece informações relevantes sobre as expectativas coletivas dos investidores, permitindo ao analista aprimorar sua tomada de decisão e reduzir os riscos associados às operações (THOMSETT, 2012).

Os indicadores técnicos compõem o principal instrumental analítico da análise técnica e consistem em cálculos matemáticos aplicados às séries históricas de preços e/ou de volume de negociação de um ativo, cujo resultado é expresso em um valor numérico utilizado para identificar padrões, tendências, condições de sobrecompra ou sobrevenda e possíveis mudanças no comportamento dos preços (ACHELIS, 2000). Esses indicadores classificam-se em dois grandes grupos: os indicadores de tendência, também denominados indicadores de atraso (*lagging indicators*), que identificam a direção predominante do mercado e são mais eficazes em períodos de tendências prolongadas; e os indicadores antecedentes (*leading indicators*), utilizados para sinalizar possíveis mudanças no comportamento dos preços, sendo mais adequados a mercados lateralizados. Um exemplo paradigmático é o *Moving Average Convergence Divergence* (MACD), calculado pela diferença entre médias móveis de 12 e 26 dias, que sinaliza alterações nas expectativas dos investidores e nas dinâmicas de oferta e demanda. Adicionalmente, os indicadores de mercado ampliam o escopo analítico ao considerarem dados agregados do mercado como um todo, categorizando-se em indicadores monetários, de sentimento e de momentum, que fornecem, respectivamente, informações sobre o ambiente econômico externo, as expectativas dos investidores e o estado atual do mercado (THOMSETT, 2012).

A seguir, são apresentados os 44 indicadores técnicos utilizados neste trabalho, organizados por categoria, juntamente com suas respectivas formulações matemáticas. Nas fórmulas, C_t, H_t, L_t, O_t e V_t denotam, respectivamente, o preço de fechamento, o máximo, o mínimo, a abertura e o volume de negociação no dia t. O parâmetro n indica o período (em dias) utilizado no cálculo de cada indicador, cujo valor padrão segue as convenções amplamente adotadas na análise técnica (ACHELIS, 2000; THOMSETT, 2012).

2.3.1 Indicadores de Volatilidade

1. **ATR²** — *Average True Range* ($n = 14$): mede a volatilidade média do ativo com base no verdadeiro intervalo de preço:

$$TR_t = \max\left(H_t - L_t, |H_t - C_{t-1}|, |L_t - C_{t-1}|\right), \tag{12}$$

$$ATR_t = \frac{1}{n}\sum_{i=0}^{n-1} TR_{t-i}\ ; \tag{13}$$

2. **BandWid** — *Bollinger Bandwidth* ($n = 20$, $k = 2$): largura relativa das Bandas de Bollinger, indicando o nível de volatilidade:

$$BandWid_t = \frac{BB_t^{sup} - BB_t^{inf}}{BB_t^{med}}, \tag{14}$$

onde $BB_t^{med} = SMA(C, n)$ é a média móvel simples do preço de fechamento, $BB_t^{sup} = BB_t^{med} + k \cdot \sigma_n$ é a banda superior, $BB_t^{inf} = BB_t^{med} - k \cdot \sigma_n$ é a banda inferior, k é o número de desvios-padrão que define a largura das bandas e σ_n é o desvio-padrão de C nos últimos n dias.

3. **Sigma** — Desvio padrão do retorno ($n = 14$): volatilidade medida pelo desvio padrão dos retornos diários:

$$Sigma_t = \sqrt{\frac{1}{n-1}\sum_{i=0}^{n-1}(r_{t-i} - \bar{r})^2}, \tag{15}$$

onde $r_t = C_t/C_{t-1} - 1$ é o retorno diário e $\bar{r}$ é a média dos retornos no período.

4. **VOLA** — *Chaikin Volatility* ($n = 10$): mede a variação percentual do spread entre máximo e mínimo:

$$VOLA_t = \frac{EMA(H - L, n)_t - EMA(H - L, n)_{t-n}}{EMA(H - L, n)_{t-n}} \times 100, \tag{16}$$

onde $EMA(H - L, n)_t$ é a média móvel exponencial da amplitude diária $(H_t - L_t)$ nos últimos n dias.

2.3.2 Indicadores de Tendência

1. **ADX** — *Average Directional Index* ($n = 14$): quantifica a força da tendência (alta ou baixa), sem indicar a direção:

$$+DI_t = \frac{EMA(+DM, n)_t}{ATR_t} \times 100, \quad -DI_t = \frac{EMA(-DM, n)_t}{ATR_t} \times 100, \tag{17}$$

$$DX_t = \frac{|+DI_t - (-DI_t)|}{+DI_t + (-DI_t)} \times 100, \quad ADX_t = EMA(DX, n)_t, \tag{18}$$

onde $+DM_t$ (*positive directional movement*) é definido como $\max(H_t - H_{t-1}, 0)$ se $H_t - H_{t-1} > L_{t-1} - L_t$, e zero caso contrário; $-DM_t$ (*negative directional movement*) é definido simetricamente; $+DI_t$ e $-DI_t$ são os indicadores direcionais positivos e negativos; e DX_t é o índice direcional antes da suavização.

2. **DIS** — *Disparity* ($n = 14$): distância percentual entre o preço e sua média móvel:

$$DIS_t = \frac{C_t - SMA(C, n)_t}{SMA(C, n)_t} \times 100, \tag{19}$$

onde $SMA(C, n)_t$ é a média móvel simples do preço de fechamento nos últimos n dias.

3. **MAO** — *Moving Average Oscillator*: diferença entre duas médias móveis exponenciais de períodos curto e longo:

$$MAO_t = EMA(C, n_1)_t - EMA(C, n_2)_t, \tag{20}$$

onde $n_1 = 5$ é o período curto e $n_2 = 35$ é o período longo.

4. **PO** — *Price Oscillator*: variação percentual entre duas médias móveis:

$$PO_t = \frac{SMA(C, n_1)_t - SMA(C, n_2)_t}{SMA(C, n_2)_t} \times 100, \tag{21}$$

onde n_1 é o período curto e n_2 é o período longo da média móvel.

5. **TRIX** — *Triple Exponential Moving Average* ($n = 12$): taxa de variação percentual de uma EMA tripla do preço de fechamento:

$$EMA3_t = EMA\big(EMA\big(EMA(C, n), n\big), n\big)_t, \tag{22}$$

$$TRIX_t = \frac{EMA3_t - EMA3_{t-1}}{EMA3_{t-1}} \times 100, \tag{23}$$

onde $EMA3_t$ é a aplicação tripla e consecutiva da média móvel exponencial sobre o preço de fechamento.

6. **MACD** — *Moving Average Convergence Divergence*: diferença entre EMA de curto e longo prazo:

$$MACD_t = EMA(C, 12)_t - EMA(C, 26)_t, \tag{24}$$

onde 12 e 26 são os períodos curto e longo, respectivamente.

7. **NBIAS** ($n = 6$): viés normalizado do preço em relação à sua média móvel:

$$NBIAS_t = \frac{C_t - SMA(C, n)_t}{SMA(C, n)_t} \; ; \tag{25}$$

8. **SMA_5 e SMA_10** — Médias Móveis Simples:

$$SMA(C, n)_t = \frac{1}{n} \sum_{i=0}^{n-1} C_{t-i}, \quad n \in \{5, 10\} \; ; \tag{26}$$

9. **EMA_5 e EMA_10** — Médias Móveis Exponenciais:

$$EMA(C, n)_t = \alpha \cdot C_t + (1 - \alpha) \cdot EMA(C, n)_{t-1}, \quad \alpha = \frac{2}{n+1}, \quad n \in \{5, 10\}, \tag{27}$$

onde α é o fator de suavização que controla o peso atribuído à observação mais recente.

2.3.3 Indicadores de Momentum

1. **WR%** — *Williams %R* ($n = 14$): posição do fechamento em relação ao intervalo de preço recente:
$$WR_t = \frac{H_n^{max} - C_t}{H_n^{max} - L_n^{min}} \times (-100), \tag{28}$$
onde $H_n^{max} = \max(H_t, \ldots, H_{t-n+1})$ é o preço máximo e $L_n^{min} = \min(L_t, \ldots, L_{t-n+1})$ é o preço mínimo observados nos últimos n dias;

2. **RSI** — *Relative Strength Index* ($n = 14$): oscilador que mede a velocidade e magnitude das variações de preço:
$$RSI_t = 100 - \frac{100}{1 + RS_t}, \quad RS_t = \frac{EMA(\Delta C^+, n)_t}{EMA(\Delta C^-, n)_t}, \tag{29}$$
onde RS_t é a força relativa, $\Delta C^+ = \max(C_t - C_{t-1}, 0)$ é a variação positiva do preço e $\Delta C^- = \max(C_{t-1} - C_t, 0)$ é a variação negativa.

3. **MOM** — *Momentum* ($n = 10$): diferença absoluta entre o preço atual e o preço de n dias atrás:
$$MOM_t = C_t - C_{t-n}; \tag{30}$$

4. **PSY** — *Psychology Line* ($n = 12$): proporção de dias de alta nos últimos n dias:
$$PSY_t = \frac{\#\{i : C_{t-i} > C_{t-i-1},\, i = 0, \ldots, n-1\}}{n} \times 100, \tag{31}$$
onde $\#\{\cdot\}$ denota a cardinalidade do conjunto, isto é, o número de dias em que o preço de fechamento foi superior ao do dia anterior;

5. **RMI** — *Relative Momentum Index* ($n = 14$, $m = 5$): variante do RSI que compara o preço com o de m dias atrás em vez do dia imediatamente anterior:
$$RMI_t = 100 - \frac{100}{1 + \dfrac{EMA(\max(C_t - C_{t-m}, 0), n)_t}{EMA(\max(C_{t-m} - C_t, 0), n)_t}}, \tag{32}$$
onde m é o período de *lookback* para o cálculo da variação de preço e n é o período de suavização da EMA;

6. **ROC** — *Rate of Change* ($n = 14$): variação percentual do preço em relação a n dias atrás:
$$ROC_t = \frac{C_t - C_{t-n}}{C_{t-n}} \times 100\ ; \tag{33}$$

7. **SROC** — *Smoothed Rate of Change* ($n = 13$): ROC aplicado a uma EMA do preço em vez do preço bruto:

$$SROC_t = \frac{EMA(C,n)_t - EMA(C,n)_{t-n}}{EMA(C,n)_{t-n}} \times 100 \; ; \tag{34}$$

8. **NCO** — *Net Change Oscillator*: variação líquida do preço de um dia para o outro:

$$NCO_t = C_t - C_{t-1}; \tag{35}$$

9. **Ret** — Retorno diário simples:

$$Ret_t = \frac{C_t - C_{t-1}}{C_{t-1}} \; ; \tag{36}$$

10. **Return** — Média móvel do retorno ($n = 14$):

$$Return_t = \frac{1}{n} \sum_{i=0}^{n-1} Ret_{t-i}, \tag{37}$$

onde Ret_{t-i} é o retorno diário definido anteriormente.

2.3.4 Indicadores Estocásticos

1. **RSV** — *Raw Stochastic Value* ($n = 14$): posição relativa do fechamento no intervalo de preço:

$$RSV_t = \frac{C_t - L_n^{min}}{H_n^{max} - L_n^{min}} \times 100, \tag{38}$$

onde H_n^{max} e L_n^{min} são, respectivamente, o preço máximo e o preço mínimo observados nos últimos n dias, conforme definidos no indicador WR%;

2. **Kvalue** — Estocástico %K ($n = 3$): média móvel do RSV, que suaviza as oscilações do valor estocástico bruto:

$$K_t = SMA(RSV,n)_t \; ; \tag{39}$$

3. **Dvalue** — Estocástico %D ($n = 3$): média móvel do %K, que atua como linha de sinal do oscilador estocástico:

$$D_t = SMA(K,n)_t \; ; \tag{40}$$

4. **Jvalue** — Estocástico %J: combinação de %K e %D que amplifica sinais de sobrecompra/sobrevenda:
$$J_t = 3K_t - 2D_t \; ; \tag{41}$$

5. **CCI** — *Commodity Channel Index* ($n = 14$): desvio do preço típico em relação à sua média:
$$TP_t = \frac{H_t + L_t + C_t}{3}, \quad CCI_t = \frac{TP_t - SMA(TP, n)_t}{0{,}015 \cdot MD_t}, \tag{42}$$
onde TP_t é o preço típico (média aritmética entre o máximo, o mínimo e o fechamento), $MD_t = \frac{1}{n}\sum_{i=0}^{n-1} |TP_{t-i} - SMA(TP, n)_t|$ é o desvio médio absoluto e a constante $0{,}015$ é um fator de escala proposto por Lambert para que aproximadamente 70% dos valores de CCI estejam entre -100 e $+100$.

2.3.5 Indicadores de Volume

1. **OBV** — *On Balance Volume*: acumula o volume nos dias de alta e subtrai nos de baixa:
$$OBV_t = OBV_{t-1} + \begin{cases} V_t, & \text{se } C_t > C_{t-1}, \\ -V_t, & \text{se } C_t < C_{t-1}, \\ 0, & \text{se } C_t = C_{t-1}, \end{cases} \tag{43}$$
onde V_t é o volume de negociação no dia t;

2. **CMF** — *Chaikin Money Flow* ($n = 20$): mede o fluxo de dinheiro ponderado pelo volume:
$$MFM_t = \frac{(C_t - L_t) - (H_t - C_t)}{H_t - L_t}, \quad CMF_t = \frac{\sum_{i=0}^{n-1} MFM_{t-i} \cdot V_{t-i}}{\sum_{i=0}^{n-1} V_{t-i}}, \tag{44}$$
onde MFM_t é o multiplicador de fluxo de dinheiro, que varia entre -1 e $+1$ conforme a posição do fechamento em relação à amplitude do dia;

3. **FI** — *Force Index* ($n = 13$): combina variação de preço e volume para medir a força dos movimentos:
$$FI_t = EMA\big((C_t - C_{t-1}) \cdot V_t,\, n\big)_t, \tag{45}$$
onde o produto $(C_t - C_{t-1}) \cdot V_t$ representa a força bruta do movimento diário, suavizada pela EMA de n períodos;

4. **MFI** — *Money Flow Index* ($n = 14$): RSI ponderado pelo volume:
$$MF_t = TP_t \cdot V_t, \quad MFI_t = 100 - \frac{100}{1 + \dfrac{\sum MF^{+}}{\sum MF^{-}}}, \tag{46}$$

onde MF_t é o fluxo de dinheiro diário, MF^+ é a soma dos fluxos em dias com $TP_t > TP_{t-1}$ e MF^- é a soma dos fluxos em dias com $TP_t < TP_{t-1}$, ambos nos últimos n dias;

5. **VMA** — *Volume Moving Average* ($n = 20$): média móvel simples do volume:

$$VMA_t = SMA(V, n)_t = \frac{1}{n} \sum_{i=0}^{n-1} V_{t-i} \; ; \tag{47}$$

6. **VOS** — *Volume Oscillator*: diferença percentual entre duas médias de volume:

$$VOS_t = \frac{SMA(V, n_1)_t - SMA(V, n_2)_t}{SMA(V, n_2)_t} \times 100, \tag{48}$$

onde $n_1 = 5$ é o período curto e $n_2 = 10$ é o período longo da média móvel do volume;

7. **VROC** — *Volume Rate of Change* ($n = 14$): taxa de variação percentual do volume:

$$VROC_t = \frac{V_t - V_{t-n}}{V_{t-n}} \times 100 \; ; \tag{49}$$

8. **AD** — *Accumulation/Distribution Line*: acumula o fluxo de dinheiro diário:

$$AD_t = AD_{t-1} + MFM_t \cdot V_t, \tag{50}$$

onde MFM_t é o multiplicador de fluxo de dinheiro definido no indicador CMF;

9. **CO** — *Chaikin Oscillator*: diferença entre EMAs rápida e lenta da linha AD:

$$CO_t = EMA(AD, 3)_t - EMA(AD, 10)_t, \tag{51}$$

onde 3 e 10 são os períodos curto e longo da média móvel exponencial aplicada à linha AD.

2.3.6 Indicadores Diversos

1. **BandPer** — *Bollinger %b* ($n = 20$, $k = 2$): posição relativa do preço dentro das Bandas de Bollinger:

$$BandPer_t = \frac{C_t - BB_t^{inf}}{BB_t^{sup} - BB_t^{inf}}, \tag{52}$$

onde BB_t^{sup} e BB_t^{inf} são as bandas superior e inferior de Bollinger, conforme definido no indicador BandWid;

2. **EOM** — *Ease of Movement* ($n = 14$): relaciona a variação de preço com o volume, normalizado pela amplitude:

$$EMV_t = \frac{(H_t + L_t)/2 - (H_{t-1} + L_{t-1})/2}{V_t/(H_t - L_t)}, \quad EOM_t = SMA(EMV, n)_t, \tag{53}$$

onde EMV_t é o valor de facilidade de movimento diário, cujo numerador mede o deslocamento do ponto médio do preço e o denominador é a razão entre o volume e a amplitude (*box ratio*);

3. **MI** — *Mass Index* ($n = 25$): detecta reversões de tendência por meio da expansão/contração do spread:

$$MI_t = \sum_{i=0}^{n-1} \frac{EMA(H - L, 9)_{t-i}}{EMA\big(EMA(H - L, 9), 9\big)_{t-i}}, \tag{54}$$

onde o numerador é a EMA simples da amplitude diária e o denominador é a EMA dupla da mesma amplitude, ambos com período de suavização igual a 9;

4. **SONAR e SONSIG**: o SONAR é um oscilador baseado na diferença entre uma EMA de curto prazo aplicada ao Williams %R e uma constante de referência. O SONSIG é a sua linha de sinal:

$$SONAR_t = EMA(WR, n_1)_t - 50, \quad SONSIG_t = EMA(SONAR, n_2)_t, \tag{55}$$

onde $n_1 = 10$ é o período de suavização do Williams %R e $n_2 = 5$ é o período da linha de sinal.

Ressalta-se que, no contexto deste trabalho, os indicadores técnicos devem ser calculados com base apenas nas informações disponíveis até o instante t, evitando a incorporação de dados futuros no processo de construção das variáveis de entrada. Esse cuidado é necessário para reduzir o risco de viés de antecipação nos modelos de aprendizado de máquina.

Além disso, em aplicações computacionais, situações em que os denominadores das expressões assumem valor nulo devem ser tratadas por meio de procedimentos específicos de pré-processamento para evitar instabilidades numéricas no cálculo dos indicadores. Neste trabalho, adotou-se uma estratégia em camadas: primeiramente, observações com preço mínimo igual a zero foram removidas antes do cálculo dos indicadores; em seguida, constantes de estabilização de magnitude $\epsilon = 10^{-10}$ foram somadas aos denominadores das expressões analíticas de cada indicador; por fim, eventuais valores ausentes remanescentes foram eliminados por remoção de linha e os casos residuais na etapa de normalização min-max foram substituídos por zero.

2.4 Considerações Finais do Capítulo

Este capítulo apresentou os principais fundamentos teóricos que sustentam a metodologia proposta nesta dissertação. Inicialmente, foram discutidos conceitos de finanças comportamentais, com destaque para a Teoria da Utilidade Esperada e a Teoria do Prospecto, cuja compreensão é necessária para fundamentar o método TODIM. Em seguida, foram apresentados os métodos de decisão multicritério e a formulação matemática do TODIM, que foi utilizado neste trabalho para apoiar a recomendação de modelos de negociação. Por fim, foram discutidos os fundamentos da análise técnica e os indicadores técnicos empregados como variáveis de entrada dos modelos de aprendizado de máquina.

3 Trabalhos Relacionados

A busca por modelos capazes de prever os movimentos do mercado de ações constitui uma trajetória que reflete a evolução das ciências de dados e da econometria. A literatura sobre previsão no mercado acionário, ampla e heterogênea, pode ser compreendida como uma progressão de paradigmas. Este capítulo revisa algumas das principais contribuições publicadas entre 2007 e 2025, sem a pretensão de esgotar o assunto, organizando-as em uma trajetória que abrange desde os modelos estatísticos clássicos, passando pelas técnicas de inteligência artificial, até as integrações mais recentes com as finanças comportamentais.

Diferentemente dos estudos que se concentram apenas na previsão de preços ou de retornos, esta dissertação propõe uma abordagem de recomendação de modelos de negociação, na qual o desempenho preditivo é avaliado em conjunto com múltiplas métricas de risco e de retorno. Além disso, enquanto parte significativa da literatura utiliza mercados internacionais, este trabalho concentra-se no mercado acionário brasileiro, considerando a organização setorial das empresas listadas na B3 e empregando modelos de *machine learning* como candidatos avaliados pelo método TODIM.

3.1 Estudos Baseados em Técnicas Estatísticas Tradicionais

A trajetória inicial da literatura sobre previsão no mercado financeiro é marcada pelo predomínio dos modelos estatísticos tradicionais, que, por muito tempo, constituíram o principal arcabouço metodológico da área. Ferramentas como ARIMA (*Autoregressive Integrated Moving Average*) e GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) forneceram bases rigorosas para a modelagem de séries temporais financeiras, sendo amplamente utilizadas para capturar tendências lineares e a dinâmica da volatilidade dos ativos. Entretanto, a dificuldade desses modelos em lidar com a não linearidade intrínseca dos mercados abriu espaço para o surgimento de abordagens alternativas.

Dando início a essa discussão, Lee, Yoo e Jin (2007) compararam o desempenho dos modelos BPNN (*Back-Propagation Neural Network*) e SARIMA (*Seasonal Autoregressive Integrated Moving Average*) na previsão do índice da bolsa de valores coreana (KOSPI), utilizando dados de 1999 a 2006. Os resultados indicaram que o modelo SARIMA apresentou maior precisão na previsão dos valores do índice, especialmente no médio prazo. Em contrapartida, o modelo BPNN apresentou melhor desempenho na previsão dos retornos do KOSPI, embora a diferença de precisão entre os dois modelos não tenha sido estatisticamente significativa.

A família de modelos ARCH e GARCH foi concebida especificamente para capturar e prever a volatilidade condicional dos dados financeiros. Em uma revisão abrangente, Engle, Focardi e Fabozzi (2008) destacam que esses modelos, juntamente com suas extensões, como TGARCH (*Threshold Generalized Autoregressive Conditional Heteroskedasticity*), EGARCH (*Exponential Generalized Autoregressive Conditional Heteroskedasticity*) e MGARCH (*Multivariate Generalized Autoregressive Conditional Heteroskedasticity*), tornaram-se instrumentos centrais da econometria financeira. Embora eficazes na modelagem da incerteza e das correlações entre ativos, esses modelos ainda enfrentam limitações significativas na previsão dos retornos esperados.

No contexto brasileiro, Faria et al. (2009) realizaram um estudo comparativo entre Redes Neurais Artificiais (RNA) e a Suavização Exponencial Adaptativa (*Adaptive Exponential Smoothing*, AES) para prever o índice Ibovespa, utilizando dados diários de 1998 a 2008. Embora ambos os modelos tenham apresentado desempenho semelhante na previsão dos valores de retorno, as RNAs mostraram-se superiores na previsão da direção do mercado, evidenciando seu potencial para o desenvolvimento de estratégias de investimento.

A previsão das tendências de cinco índices do mercado acionário indiano foi investigada por Merh, Saxena e Pardasani (2010) por meio de dois modelos híbridos que combinam Redes Neurais Artificiais e ARIMA. Os resultados indicaram que a eficácia das abordagens dependeu do índice analisado: o modelo RNA_ARIMA apresentou melhor desempenho nos índices SENSEX, BSE IT e SP CNX Nifty, enquanto o ARIMA_RNA foi ligeiramente superior no índice BSE 100. Para o altamente volátil índice BSE Oil & Gas, apenas os modelos baseados em RNA conseguiram gerar previsões viáveis, uma vez que o ARIMA não lidou adequadamente com a heteroscedasticidade dos dados.

A metodologia de Box-Jenkins foi aplicada por Adebiyi, Adewumi e Ayo (2014) para a construção de modelos ARIMA voltados à previsão de preços de ações, utilizando dados das bolsas de valores de Nova York e da Nigéria. O processo envolveu testes de estacionariedade, como o ADF (*Augmented Dickey-Fuller*), e o uso do critério BIC (*Bayesian Information Criterion*) para a seleção do modelo. Os autores concluíram que o ARIMA apresenta bom desempenho em previsões de curto prazo, podendo servir de suporte à tomada de decisão dos investidores.

De forma semelhante, Idrees, Alam e Agarwal (2019) empregaram o modelo ARIMA para prever a volatilidade dos índices indianos Nifty e Sensex, com dados coletados entre 2012 e 2016. A abordagem incluiu testes de estacionariedade e a identificação dos parâmetros do modelo por meio de gráficos de autocorrelação e de autocorrelação parcial. O modelo ARIMA(0,1,0), após diferenciação, foi utilizado para prever os valores de 2016, apresentando, segundo os autores, um desvio médio satisfatório em relação aos dados reais.

A eficácia do modelo SARIMA na previsão dos preços das ações da Apple foi inves-

tigada por Daryl et al. (2021). Após testar diferentes configurações, os autores concluíram que o modelo não conseguiu produzir previsões precisas. Essa limitação foi atribuída à elevada volatilidade do mercado acionário, e não a falhas inerentes ao modelo estatístico.

Mais recentemente, Ievsieieva et al. (2024) analisaram a aplicação conjunta dos modelos ARIMA, VAR (*Vector Autoregression*) e GARCH para a previsão de variáveis financeiras relevantes para a gestão corporativa. Os resultados indicaram que nenhum dos modelos é universalmente superior: o ARIMA mostrou-se eficaz na previsão de retornos, o VAR capturou adequadamente a interdependência entre variáveis, e o GARCH, embora menos preciso, destacou-se na análise de risco ao modelar a instabilidade do mercado. Os autores recomendam o uso combinado dessas abordagens para uma visão mais abrangente no processo decisório.

Em síntese, os estudos baseados em métodos estatísticos tradicionais confirmam sua relevância como pilares da econometria financeira. Modelos da família GARCH consolidaram-se como ferramentas essenciais para a modelagem da volatilidade, enquanto o ARIMA permanece amplamente validado para previsões de curto prazo. No entanto, a própria literatura evidencia limitações importantes: os resultados são frequentemente heterogêneos e, em mercados caracterizados por elevada volatilidade e heteroscedasticidade, modelos como o SARIMA podem apresentar desempenho insatisfatório. Essas limitações, associadas a evidências de que as técnicas de inteligência artificial tendem a ser mais eficazes na previsão da direção do mercado, indicam que os métodos estatísticos clássicos, embora fundamentais, podem ser insuficientes quando utilizados isoladamente.

3.2 Estudos Baseados em Técnicas de Inteligência Artificial

A revolução metodológica na previsão do mercado financeiro foi impulsionada pelo avanço das técnicas de inteligência artificial. O desenvolvimento de modelos como Redes Neurais Artificiais (RNA; do inglês *Artificial Neural Networks*, ANN), *Support Vector Machines* (SVM) e, mais recentemente, arquiteturas de aprendizado profundo, como *Long Short-Term Memory* (LSTM), possibilitou a modelagem de padrões complexos e de dependências temporais de longo prazo, antes difíceis de capturar. Essa vertente de pesquisa, amplamente explorada na literatura recente em previsão financeira, tem como foco principal aprimorar a capacidade preditiva dos modelos, seja na estimação de preços e retornos, seja na previsão da direção dos movimentos de mercado ou na geração de sinais de negociação.

Iniciando essa linha de investigação, Dutta, Bandopadhyay e Sengupta (2012) utilizaram o modelo de Regressão Logística (do inglês *Logistic Regression*, LR) para prever o desempenho das ações no mercado indiano, classificando-as como “boas” ou “ruins” em relação ao índice NIFTY. O modelo foi aplicado a um conjunto de dados composto por 30 grandes empresas no período de 2005 a 2008, e os autores sugerem que a abordagem pode

ser uma ferramenta prática para auxiliar investidores na seleção de ações com desempenho superior.

Por sua vez, Barak e Modarres (2015) propuseram um modelo em três etapas para prever risco e retorno de ações, combinando técnicas de mineração de dados e algoritmos de classificação, como *Decision Tree* (DT), *Decision Table + Naive Bayes* (DTNB), *k-Nearest Neighbors* (kNN), SVM e *Multilayer Perceptron* (MLP). A metodologia incluiu a identificação de variáveis influentes, a aplicação de classificadores e um algoritmo híbrido de seleção de características que combinava nove métodos distintos. Testado com dados da Bolsa de Valores de Teerã entre 2003 e 2012, o estudo demonstrou que a seleção híbrida de atributos melhorou significativamente a precisão das previsões de risco e retorno.

A integração de dados sociais com métodos tradicionais de análise também se mostrou promissora. Attigeri et al. (2015) propuseram uma abordagem que combina análise técnica e fundamental, utilizando Regressão Logística sobre dados históricos de preços e, simultaneamente, extraindo sentimentos de mídias sociais e notícias por meio de ferramentas como o *Hadoop*. Os resultados destacaram o impacto significativo do sentimento social no comportamento do mercado.

O uso de Redes Neurais Artificiais para prever os retornos do índice Nikkei 225 foi investigado por Qiu, Song e Akagi (2016). O estudo envolveu a seleção de 18 variáveis relevantes por meio de superfícies *fuzzy* e a otimização da rede neural com algoritmos metaheurísticos, como *Genetic Algorithm* (GA) e *Simulated Annealing* (SA), visando superar as limitações do método tradicional de *backpropagation*. Os resultados indicaram que os modelos híbridos, especialmente o GA-BP, apresentaram maior precisão preditiva e menor erro.

Uma abordagem híbrida, combinando Análise de Componentes Independentes (do inglês *Independent Component Analysis*, ICA) para extração de características e métodos de kernel, como SVM, para classificação, foi proposta por Ince e Trafalis (2017). Avaliado nos índices Dow Jones, Nasdaq e S&P 500, o modelo apresentou maior precisão preditiva e maior retorno sobre o investimento quando comparado a modelos individuais, demonstrando robustez na captura da complexidade das séries temporais financeiras.

A comparação entre diferentes arquiteturas de aprendizado profundo foi explorada por Hiransha et al. (2018), que avaliaram MLP, Redes Neurais Recorrentes (RNN), LSTM e Redes Neurais Convolucionais (CNN) para a previsão de preços de ações. Utilizando dados diários de fechamento da bolsa indiana (NSE) e testando em empresas da NSE e da NYSE, os autores observaram que o modelo CNN apresentou o melhor desempenho, superando inclusive o ARIMA, ao capturar dinâmicas não lineares e mudanças abruptas nos dados.

Abordagens mais recentes passaram a explorar modelos generativos. Zhang et al. (2019) propuseram uma arquitetura baseada em Redes Generativas Adversárias (GANs),

utilizando uma LSTM como gerador e um MLP como discriminador, para prever preços de fechamento diários de ações do índice S&P 500. Os resultados indicaram que o modelo GAN superou tanto o SVR quanto o LSTM isolado, apresentando erros menores e um retorno médio diário superior.

Diversas extensões do LSTM também foram propostas. Pang et al. (2020) apresentaram os modelos ELSTM e AELSTM, aplicados à previsão de índices chineses, enquanto Nabipour et al. (2020) demonstraram a superioridade do LSTM em relação a outros métodos de aprendizado de máquina para prever preços de ações em diferentes setores da Bolsa de Teerã. Estudos subsequentes reforçaram esse desempenho, destacando o LSTM como uma ferramenta eficaz para capturar padrões temporais complexos (CHHAJER; SHAH; KSHIRSAGAR, 2022; SAXENA; SHARMA; GUPTA, 2023; GUPTA; JAISWAL, 2024).

A hibridização de modelos tornou-se uma tendência recorrente na literatura recente. Abordagens como GARCH-LSTM (PAN; TANG; WANG, 2024; MUTINDA; LANGAT, 2024), GAN-CNN-LSTM (VULLAM et al., 2023), SVMD-LSTM (AGARWAL et al., 2025) e MEMD-AO-LSTM (GE, 2025) têm relatado ganhos de desempenho ao combinar técnicas estatísticas, métodos de decomposição de sinais e redes neurais profundas. Paralelamente, o uso de dados não estruturados, como sentimentos extraídos de mídias sociais, mostrou-se relevante para aprimorar a capacidade preditiva (MEHTA; PANDYA; KOTECHA, 2021; ZAICHENKO et al., 2023).

Arquiteturas de aprendizado por reforço (YASIN; PASCHKE; QUNDUS, 2024), modelos *neuro-fuzzy* (SAFARI; GHAEMI, 2025) e sistemas baseados em *Transformers* também ampliaram a fronteira da pesquisa, oferecendo soluções robustas em cenários de alta volatilidade e com dados financeiros não estacionários.

Em síntese, os estudos baseados em técnicas de inteligência artificial evidenciam uma clara evolução rumo a modelos híbridos e arquiteturas profundas. Redes LSTM e suas variantes consolidaram-se como ferramentas centrais devido à capacidade de modelar dependências temporais de longo prazo. Os trabalhos mais recentes demonstram que os melhores resultados são obtidos pela integração dessas redes com modelos estatísticos clássicos, técnicas de decomposição de sinais, metaheurísticas de otimização e fontes alternativas de informação, como dados de sentimento. Essa convergência tem apresentado, em diversos estudos, desempenho superior ao dos métodos tradicionais, embora os resultados dependam do mercado analisado, do período considerado, das métricas utilizadas e da estratégia de validação adotada.

Apesar dos avanços observados, é importante destacar que a melhora nas métricas preditivas não implica, necessariamente, melhor desempenho econômico em estratégias de negociação. A previsão precisa de preços, retornos ou da direção do mercado constitui apenas uma etapa do processo decisório. Para que um modelo seja efetivamente útil ao investidor, é necessário examinar também sua capacidade de gerar sinais operaciona-

lizáveis, sua exposição ao risco, sua estabilidade na generalização e seu desempenho sob condições realistas de mercado, incluindo custos de transação, liquidez e frequência de negociação. Assim, a literatura recente evidencia a necessidade de complementar a avaliação preditiva com critérios financeiros, psicológicos e operacionais, aproximando os modelos de previsão dos sistemas efetivos de recomendação e tomada de decisão.

3.3 Estudos Baseados em Técnicas Multicritério e Teoria do Prospecto

Nesse contexto, prever preços com alta precisão representa apenas uma parte do desafio. A construção de um sistema de recomendação eficaz exige não apenas modelos preditivos robustos, mas também mecanismos capazes de representar o processo de decisão de investimento sob risco. É nesse cenário que emerge uma terceira corrente de pesquisa, fundamentada em métodos multicritério e na Teoria do Prospecto, que incorpora aspectos comportamentais do investidor ao processo decisório. Ao considerar a aversão à perda e a percepção assimétrica de ganhos e perdas, esses modelos buscam desenvolver estratégias de negociação mais alinhadas ao comportamento real dos agentes financeiros.

A incorporação de medidas oriundas da Teoria do Prospecto em modelos preditivos foi investigada por Zhang (2024) por meio da construção de um "índice prospectivo". Esse índice foi utilizado como variável explicativa em modelos de regressão linear e de *Random Forest* para prever o S&P 500 e o *Bitcoin*. Os resultados indicaram que a inclusão do índice teve efeito marginal nas previsões do S&P 500, mas melhorou significativamente as métricas do *Bitcoin*, sugerindo que sinais comportamentais tendem a ser mais informativos em mercados potencialmente menos eficientes. O autor também destacou limitações, como o risco de *overfitting*.

Por sua vez, Puppo et al. (2022) apresentaram um sistema automático de seleção de regras de negociação que combina um grande conjunto de regras técnicas com o método multicritério TODIM, inspirado na Teoria do Prospecto. O TODIM é utilizado para ponderar simultaneamente retorno e múltiplas medidas de risco, incorporando a aversão à perda para calcular a utilidade relativa de cada regra. Em experimentos com oito ações brasileiras, o procedimento demonstrou reduzir a exposição ao risco e melhorar as medidas ajustadas ao risco, em comparação com abordagens alternativas, embora as diferenças na rentabilidade bruta nem sempre tenham sido estatisticamente significativas. O estudo conclui que a incorporação das assimetrias de ganhos e perdas, por meio do TODIM, contribui para uma seleção de regras mais consistente sob uma perspectiva multicritério.

A combinação de variáveis *fuzzy* (para representar incerteza), da teoria das três decisões (*Three-way Decision*, 3WD) e da *Cumulative Prospect Theory* (CPT) fundamentou o modelo de seleção de portfólio proposto por Wang et al. (2022). Nesse modelo, a 3WD

é utilizada para classificar ativos em “comprar”, “esperar” ou “vender”, enquanto a CPT representa preferências comportamentais do investidor. Na prática, a abordagem calcula valores prospectivos (CPT), classifica os ativos (3WD) e, em seguida, executa a etapa de otimização do portfólio. Testes no Dow Jones e na NYSE indicaram redução de risco e melhora no desempenho ajustado ao risco, resultando em uma alocação mais conservadora.

Ao analisar ativos por setor, Lv et al. (2023) propuseram um algoritmo de recomendação de modelos de *trading* baseado no método TODIM para explorar a similaridade de comportamento entre ações do mesmo segmento. Os autores dividiram as ações do S&P 500 e do CSI 300 em nove indústrias, extraíram 44 indicadores técnicos, utilizaram a análise *walk-forward* para gerar sinais diários de negociação (compra no fechamento e venda no fechamento do dia seguinte) e realizaram *backtests* com base em oito métricas de desempenho. Como candidatos, foram avaliados doze algoritmos (métodos tradicionais e redes profundas), e foi construída uma matriz de decisão com pesos atribuídos a cada métrica. O TODIM, então, realiza comparações pareadas e gera um ranqueamento, recomendando o melhor modelo para cada indústria. Além disso, foram propostas métricas específicas de avaliação de generalização, validadas em conjuntos *in-sample* e *out-of-sample*, indicando que a abordagem recomenda modelos com capacidade de generalização adequada e permite a construção de carteiras *long–short* com retornos superiores ao *benchmark* em diferentes cenários.

Por último, Jain, Sahay e Nupur (2025) sugeriram uma estrutura híbrida em duas etapas para a seleção e a otimização de carteiras. Na primeira etapa, o modelo utiliza *Robust Expectation Maximization* (REM) para agrupar ativos e aplica o método TODIM para ranquear e selecionar empresas nesses grupos. Na segunda etapa, os retornos são estimados por um modelo conjunto (*N-BEATS–XGBoost*) e formula-se um problema de otimização multiobjetivo, incorporando medidas assimétricas de risco, momentos superiores e critérios de diversificação, com a possibilidade de incluir restrições ESG (*Environmental, Social and Governance*). Avaliado no índice Nifty-100 (2018–2023), a estrutura proposta apresentou ganhos em diversificação, controle de risco e alinhamento sustentável em comparação com abordagens clássicas, como Média–Variância.

Em síntese, a incorporação da Teoria do Prospecto, especialmente por meio de métodos multicritério, como o TODIM, representa uma evolução relevante na modelagem de decisões de investimento, ao transcender a mera maximização do retorno e contemplar a psicologia do investidor. Os estudos indicam que considerar a aversão à perda e as assimetrias na percepção de risco pode levar a estratégias mais robustas e conservadoras, com redução da exposição ao risco e melhoria das métricas ajustadas ao risco. Além disso, a aplicação dessas técnicas na recomendação setorial de modelos de negociação e sua integração em estruturas híbridas de seleção de portfólio evidenciam a versatilidade dessa abordagem. Tais métodos mostram-se úteis tanto para buscar desempenho superior

a índices de referência quanto para promover alocações mais diversificadas e, quando aplicável, alinhadas a critérios de sustentabilidade.

3.4 Considerações e Resumo dos Trabalhos

A revisão da literatura sobre previsão e negociação no mercado de ações, contemplando trabalhos publicados aproximadamente entre 2007 e 2025, revela uma trajetória evolutiva bem definida nas abordagens metodológicas adotadas pela comunidade científica. O ponto de partida, amplamente documentado, é a transição gradual dos modelos estatísticos clássicos para arquiteturas baseadas em aprendizado de máquina e de aprendizado profundo, as quais se mostraram mais adequadas para capturar a complexidade, a não linearidade e a dinâmica temporal dos dados financeiros.

Os trabalhos analisados evidenciam que a sofisticação dos modelos preditivos avançou significativamente ao longo do tempo. A incorporação de fontes de dados alternativas, como o sentimento extraído de mídias sociais e notícias financeiras, bem como o desenvolvimento de modelos híbridos que combinam o poder preditivo de redes neurais recorrentes (como LSTM e suas variantes) e arquiteturas generativas (como GANs) com técnicas de seleção e extração de características, consolidou-se como uma prática recorrente. Ademais, a integração de modelos econométricos tradicionais, como o GARCH, com redes neurais profundas destacou-se como uma estratégia robusta para lidar simultaneamente com a volatilidade condicional e com as dependências temporais de curto e de longo prazo presentes nos ativos financeiros.

Entretanto, esta revisão também evidencia o surgimento e a consolidação de uma terceira vertente de pesquisa, que avança além da mera previsão de preços e passa a tratar explicitamente da complexidade do processo de tomada de decisão sob risco. Em vez de focar exclusivamente na estimação do valor futuro de um ativo, esses estudos concentram-se no desenvolvimento de sistemas de recomendação, seleção de ativos e construção de portfólios que incorporem, de forma explícita, os vieses comportamentais do investidor. Nesse contexto, a aplicação de conceitos oriundos da Teoria do Prospecto assume um papel central, ao reconhecer que decisões financeiras reais são guiadas por uma percepção assimétrica de ganhos e perdas e por uma racionalidade limitada.

Trabalhos recentes demonstram que a incorporação de funções de valor prospectivo e a utilização de métodos multicritério inspirados nessa teoria, como o método TODIM, possibilitam a construção de estratégias de negociação mais robustas. Diferentemente das abordagens tradicionais, cujo objetivo principal é maximizar o retorno esperado, esses modelos buscam otimizar métricas ajustadas ao risco e promover a construção de carteiras mais resilientes, coerentes com o perfil psicológico e a aversão à perda do tomador de decisão.

Em síntese, a literatura aponta claramente para uma fronteira de pesquisa caracterizada pela convergência entre o elevado poder preditivo da inteligência artificial e o realismo comportamental proporcionado pela Teoria do Prospecto. Uma tendência relevante na literatura recente consiste em não limitar a análise à previsão do mercado, mas sim utilizar essas previsões como insumo para sistemas de decisão capazes de incorporar múltiplos critérios e preferências do investidor. É precisamente nessa interseção que se insere a presente dissertação, ao propor um modelo que integra técnicas de *machine learning* à tomada de decisão multicritério por meio do método TODIM, com foco no mercado acionário brasileiro.

Nesse contexto, o presente trabalho utiliza como referência o artigo de Lv et al. (2023), apresentando três diferenças principais. A primeira é a adaptação do arcabouço original, aplicado aos mercados americano e chinês, ao mercado acionário brasileiro (B3). A segunda é a inclusão de uma etapa de seleção de atributos, ausente no estudo original, que combina a correlação de Pearson, para eliminar a redundância entre os indicadores técnicos, e a informação mútua, para filtrar os atributos mais relevantes em relação à variável-alvo. A terceira e principal diferença é a extensão da granularidade da recomendação: enquanto Lv et al. (2023) recomendam um modelo ótimo por setor, esta dissertação propõe também uma recomendação em nível de ação individual, ampliando o escopo original que se limitava à granularidade setorial.

Por último, julgou-se pertinente compilar e sistematizar as principais características dos trabalhos analisados. Essa síntese tem como objetivo servir tanto como material de apoio à continuidade desta pesquisa quanto como referência para estudos futuros sobre o tema. Como resultado, o Apêndice A apresenta um resumo estruturado dos 53 artigos considerados nesta revisão.

4 Abordagem Proposta

Este capítulo apresenta a abordagem proposta para recomendação de modelos de *trading* em um mercado de ações. A metodologia, ilustrada no fluxograma da Figura 3, organiza-se em quatro etapas sequenciais: (i) obtenção e classificação setorial dos dados, (ii) cálculo e seleção de indicadores técnicos, (iii) treinamento dos modelos candidatos por meio de *Walk-Forward Analysis* e avaliação por *backtesting*, e (iv) recomendação do modelo por meio do método TODIM (Tomada de Decisão Interativa Multicritério). Avaliam-se duas granularidades de recomendação: por setor econômico e por ação individual. A seguir, cada etapa é resumida.

A primeira etapa, **Construção do Conjunto de Dados** (Seção 4.1), parte dos dados diários OHLCV de cada ação, em um período definido, e os transforma em conjuntos de dados adequados ao treinamento de modelos de aprendizado de máquina. Para cada ação, calculam-se 44 indicadores técnicos, define-se um rótulo binário indicando a direção do retorno no dia seguinte e aplicam-se normalização e, opcionalmente, seleção de atributos via correlação de Pearson e Informação Mútua. Ao final, cada ação possui duas versões de *dataset*: uma com os 44 indicadores e outra com o subconjunto selecionado.

A segunda etapa, **Execução dos Modelos de Aprendizado de Máquina** (Seção 4.2), treina cada um dos 12 modelos candidatos utilizando ambas as versões de *dataset* por meio da metodologia *Walk-Forward Analysis* (WFA). Em cada rodada, os modelos são treinados em janelas deslizantes de 250 dias e avaliados nos 5 dias subsequentes, gerando uma série temporal de sinais binários de negociação para cada ação. O WFA é executado de forma independente para cada ação e para cada modelo.

A terceira etapa, ***Backtesting* e Cálculo das Métricas de Desempenho** (Seção 4.3), aplica os sinais de negociação aos preços de fechamento para simular a execução de uma estratégia *long-only*. O retorno diário do modelo é obtido multiplicando o retorno diário do ativo pelo sinal emitido no dia anterior. A partir dessa série de retornos, calculam-se oito métricas de desempenho (WR, ARR, ASR, MDD, CAR, SOR, MSR e ART), que abrangem as dimensões de acurácia, retorno e risco. Ao final, para cada ação e para cada modelo, obtém-se um vetor de oito métricas, calculado em ambas as configurações de *features*.

A quarta etapa, **Recomendação do Modelo** (Seção 4.4), utiliza as métricas obtidas como entrada para o TODIM. As oito métricas compõem uma matriz de decisão em que as linhas representam os 12 modelos candidatos e as colunas, os critérios de avaliação. O TODIM compara os modelos par a par, incorporando aversão à perda por meio da teoria do prospecto, e produz um ranking de recomendação. Essa recomendação é realizada em duas granularidades: (i) setorial, agregando-se as métricas das ações de cada setor

pela média aritmética, e (ii) individual, utilizando-se diretamente as métricas de cada ação. A comparação entre as duas granularidades permite avaliar se a agregação setorial é suficiente ou se a heterogeneidade intrassetorial justifica recomendações individualizadas.

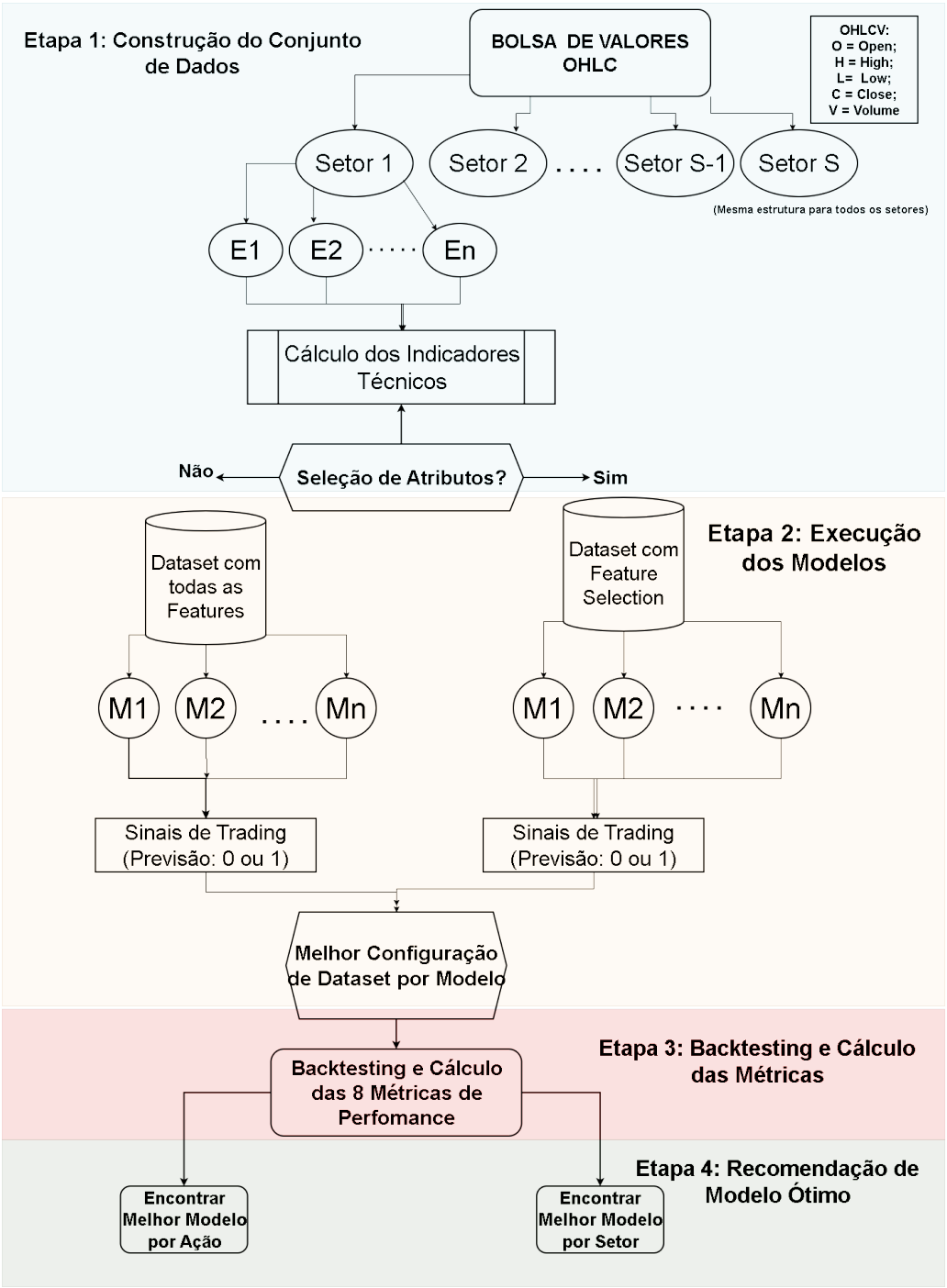

Figura 3 – Fluxograma da abordagem proposta para recomendação de modelos de *trading*.

4.1 Construção do Conjunto de Dados

A primeira etapa da abordagem proposta consiste na construção dos conjuntos de dados utilizados como entrada nos modelos de aprendizado de máquina. Esse processo é executado de forma independente para cada ação da amostra e é composto por quatro passos sequenciais: (i) obtenção dos dados OHLCV, (ii) cálculo dos indicadores técnicos, (iii) definição do rótulo binário e (iv) normalização e seleção opcional de atributos.

O primeiro passo consiste em obter os dados históricos diários de negociação da bolsa de valores analisada. Para cada ação, coletam-se o preço de abertura (*Open*, O_t), o preço máximo (*High*, H_t), o preço mínimo (*Low*, L_t), o preço de fechamento (*Close*, C_t) e o volume negociado (*Volume*, V_t), que compõem o conjunto OHLCV. Os dados abrangem um período predefinido; adicionalmente, o primeiro ano da série é reservado exclusivamente para viabilizar o cálculo dos indicadores técnicos, uma vez que diversos indicadores dependem de janelas móveis.

O segundo passo consiste no cálculo de 44 indicadores técnicos a partir dos dados OHLCV de cada ação. Esses indicadores descrevem diferentes aspectos da dinâmica de preços e volumes, abrangendo as categorias de volatilidade, tendência, *momentum*, estocásticos, de volume e diversos (Tabela 1). Ressalta-se que as variáveis brutas OHLCV não são utilizadas diretamente como *features*; são empregadas como insumo para a construção de variáveis derivadas, representadas por indicadores. As formulações matemáticas e as descrições de cada indicador são apresentadas na Seção 2.3.

Tabela 1 – Indicadores técnicos utilizados como variáveis de entrada, com tipo e quantidade por categoria.

Tipo	**Quant.**	**Indicadores**
Volatilidade	04	ATR, BandWid, Sigma, VOLA
Tendência	11	ADX, DIS, MAO, PO, TRIX, MACD, NBIAS, SMA_5, SMA_10, EMA_5, EMA_10
Momentum	10	WR, RSI, MOM, PSY, RMI, ROC, SROC, NCO, Ret, Return
Estocásticos	05	RSV, Kvalue, Dvalue, Jvalue, CCI
Volume	09	OBV, CMF, FI, MFI, VMA, VOS, VROC, AD, CO
Diversos	05	BandPer, EOM, MI, SONAR, SONSIG
Total	**44**	

O terceiro passo consiste na definição do rótulo binário (*label*), associado à variável de saída da tarefa de classificação. Para cada dia t, o rótulo é determinado pelo sinal do retorno logarítmico no dia seguinte:

$$\text{Label}_t = \begin{cases} 1, & \text{se } \ln\left(\frac{C_{t+1}}{C_t}\right) \geq 0 \\ 0, & \text{caso contrário} \end{cases}, \tag{56}$$

em que C_t representa o preço de fechamento no dia t. Dessa forma, Label$_t = 1$ indica retorno não negativo no dia seguinte, enquanto Label$_t = 0$ indica retorno negativo. Com os indicadores e o rótulo definidos, constrói-se o *dataset* da ação como uma matriz com T linhas (dias de negociação) e 46 colunas (data, 44 *features* e rótulo).

O quarto passo consiste em (i) normalizar as variáveis de entrada e (ii) aplicar, opcionalmente, seleção de atributos, gerando uma versão alternativa do *dataset* com dimensionalidade reduzida. Dessa forma, para cada ação são construídas e avaliadas duas configurações de entrada: uma com todas as 44 *features* (indicadores técnicos) normalizadas via Min–Max, e outra contendo apenas um subconjunto selecionado, também normalizado via Min–Max. Assim, além de se comparar o desempenho entre modelos de aprendizado, avalia-se também, para cada ação e para cada setor, se utilizar o conjunto completo de indicadores proporciona melhores resultados ou se a redução de dimensionalidade por meio de seleção de atributos é mais vantajosa.

A seleção de atributos é realizada em dois estágios. No primeiro estágio, calcula-se a correlação de Pearson entre pares de indicadores e, sempre que dois indicadores apresentam correlação absoluta superior a um limiar pré-definido ($|\rho| > 0{,}85$), remove-se o indicador com menor Informação Mútua em relação ao rótulo, eliminando redundâncias. No segundo estágio, calcula-se a Informação Mútua entre cada indicador remanescente e o rótulo, selecionando-se, então, os *top*-k indicadores mais informativos. A fundamentação metodológica dessas técnicas pode ser consultada na Seção 5.1.2.

Ao final do processo, obtêm-se, para cada ação, duas versões de *dataset*:

- ***Dataset* completo:** contém os 44 indicadores técnicos normalizados por Min–Max e o rótulo binário;
- ***Dataset* com seleção:** contém apenas os k indicadores selecionados pelos procedimentos de Correlação de Pearson e Informação Mútua, com as *features* normalizadas por Min–Max, permitindo avaliar o impacto da redução de dimensionalidade no desempenho dos modelos.

Ambas as versões são empregadas no treinamento e na avaliação dos 12 modelos candidatos na etapa seguinte, permitindo comparar, de forma controlada, as duas configurações de *features* para cada modelo, ação e setor.

4.2 Execução dos Modelos de Aprendizado de Máquina

Para cada ação, as duas configurações de *dataset* descritas na Seção 4.1 (conjunto completo de indicadores e conjunto com seleção de atributos), ambas com normalização Min–Max, são utilizadas como entrada para o treinamento e a avaliação de 12 modelos candidatos de aprendizado de máquina. Esses modelos foram selecionados para abranger

uma ampla variedade de paradigmas de aprendizado, garantindo que o algoritmo de recomendação TODIM avalie candidatos com estruturas fundamentalmente distintas. Os modelos são divididos em dois grupos.

O primeiro grupo compreende seis modelos tradicionais. A Regressão Logística (LR) é um modelo linear que estima a probabilidade de classe por meio de uma função sigmoide, sendo um dos modelos mais simples e interpretáveis do conjunto (ZHOU, 2021). O *Support Vector Machine* (SVM) com *kernel* RBF (*Radial Basis Function*) mapeia os dados para um espaço de alta dimensão, no qual busca um hiperplano de separação ótima, sendo capaz de capturar relações não lineares (ZHOU, 2021). O *Naive Bayes* (NB) é um classificador probabilístico que assume independência condicional entre as *features*, o que o torna particularmente sensível à presença de variáveis correlacionadas (ZHOU, 2021; LANIADO et al., 2016). A Árvore de Decisão (CART) realiza partições recursivas do espaço de *features* com base no critério de Gini, aplicando uma forma implícita de seleção de variáveis em cada nó (ZHOU, 2021). O *Random Forest* (RF) é um *ensemble* de 500 árvores de decisão treinadas com subconjuntos aleatórios de *features* e de amostras, reduzindo a variância por meio de agregação (*bagging*) (CHEN; GUESTRIN, 2016). O *XGBoost* (XGB) é um conjunto (*ensemble*) baseado em impulsionamento (*boosting*) com 15 rodadas de aprendizado sequencial, em que cada árvore corrige os erros das anteriores (CHEN; GUESTRIN, 2016).

O segundo grupo compreende seis modelos de aprendizado profundo. O *Multilayer Perceptron* (MLP) é uma rede neural densa com arquitetura de 25–15–10–5 neurônios e ativação sigmoide, sem mecanismo de recorrência (MURPHY, 2012). A *Deep Belief Network* (DBN) utiliza pré-treinamento não supervisionado antes do ajuste fino supervisionado, aprendendo representações hierárquicas dos dados (GOODFELLOW; BENGIO; COURVILLE, 2016). O *Stacked Autoencoder* (SAE) segue um princípio semelhante ao DBN, com camadas de compressão 20–10–5 que reduzem a dimensionalidade das *features* antes da classificação (MURPHY, 2012). A *Recurrent Neural Network* (RNN) é uma rede recorrente simples que processa sequências temporais, sem mecanismos de *gates* (GOODFELLOW; BENGIO; COURVILLE, 2016). O *Long Short-Term Memory* (LSTM) estende a RNN com três *gates* (*forget*, *input* e *output*) que controlam o fluxo de informação ao longo do tempo, permitindo capturar dependências de longo prazo (GOODFELLOW; BENGIO; COURVILLE, 2016). O *Gated Recurrent Unit* (GRU) é uma variante simplificada do LSTM, com dois *gates* (*reset* e *update*), oferecendo capacidade similar a menor custo computacional (GOODFELLOW; BENGIO; COURVILLE, 2016). Exceto pelo MLP, os demais modelos de aprendizado profundo utilizam a arquitetura empilhada de 20–10–5 neurônios nas camadas ocultas (GOODFELLOW; BENGIO; COURVILLE, 2016).

A diversidade arquitetural dos 12 modelos é intencional: modelos lineares (LR), baseados em *kernel* (SVM), probabilísticos (NB), baseados em árvores (CART, RF, XGB),

redes densas (MLP), redes com pré-treinamento (DBN, SAE) e redes recorrentes com e sem *gates* (RNN, LSTM, GRU) representam paradigmas distintos, cada um com vantagens e limitações específicas para a tarefa de previsão de séries financeiras. Essa variedade justifica o uso de um método de decisão multicritério para identificar qual paradigma é mais adequado para cada ação e, posteriormente, para cada setor e ação individual.

Os modelos são treinados e avaliados por meio da metodologia *Walk-Forward Analysis* (WFA), que simula o uso prático ao longo do tempo. Em cada rodada, o modelo é treinado com uma janela fixa de $d_1 = 250$ dias e, em seguida, gera previsões para os $d_2 = 5$ dias subsequentes. A janela desliza por d_2 dias e o processo se repete, conforme ilustrado na Figura 4.

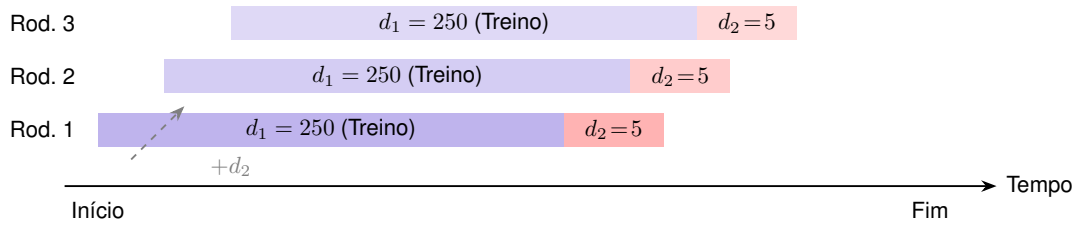

Figura 4 – Método WFA para treinamento e previsão.

Ao final do WFA, cada modelo produz uma série temporal de sinais binários: 1 (comprar a ação no fechamento do dia e vender no fechamento do dia seguinte) ou 0 (não operar). Essa abordagem constitui uma estratégia de *trading long-only* com período de detenção de 1 dia.

Para formalizar o procedimento, considere que o *dataset* D_i da i-ésima ação possui L observações diárias e que N é o comprimento da lista de símbolos das ações. O número total de rodadas do WFA é determinado pela divisão inteira entre o espaço disponível após a primeira janela de treino e o tamanho da janela de teste:

$$Q = \left\lfloor \frac{L - d_1}{d_2} \right\rfloor, \tag{57}$$

onde o operador $\lfloor \cdot \rfloor$ denota o arredondamento para baixo ao inteiro mais próximo. Como a divisão $(L - d_1)/d_2$ nem sempre resulta em um número inteiro, pode haver um resíduo de observações, denotado por H, calculado como:

$$H = (L - d_1) - d_2 \cdot Q. \tag{58}$$

As primeiras H observações do *dataset* são descartadas, de modo que as $d_1 + Q \cdot d_2$ observações restantes se encaixem perfeitamente na estrutura do WFA. A título de exemplo, para uma ação com $L = 869$ observações, obtém-se $Q = \lfloor 619/5 \rfloor = 123$ rodadas e $H = 619 - 615 = 4$ observações a serem descartadas, visto que $(L - d_1)/d_2 = 619/5 = 123{,}8$.

Após o alinhamento, o procedimento itera sobre as Q rodadas. Em cada rodada j, o conjunto de treinamento compreende as observações da posição $1 + d_2(j-1)$ até $d_1 + d_2(j-1)$, totalizando sempre $d_1 = 250$ dias. O conjunto de teste é formado pelas $d_2 = 5$ observações imediatamente subsequentes, da posição $d_1 + d_2(j-1) + 1$ até $d_1 + d_2 \cdot j$. Na primeira rodada ($j = 1$), o treino utiliza as observações de 1 a 250 e o teste, as de 251 a 255. Na segunda rodada ($j = 2$), a janela desliza d_2 posições: o treino utiliza as observações de 6 a 255 e o teste, as de 256 a 260, e assim sucessivamente.

Um aspecto importante do procedimento é a reinicialização dos parâmetros do modelo a cada rodada, isto é, em cada iteração, o modelo é treinado do zero, sem reter informações aprendidas em rodadas anteriores. Essa escolha metodológica garante que cada rodada constitui um experimento independente, evitando que o modelo acumule viés proveniente de períodos anteriores.

Ao final de cada rodada, o modelo treinado é aplicado ao conjunto de teste para gerar probabilidades preditivas $P(t)$ para cada dia do período de teste. Essas probabilidades são convertidas em sinais binários por meio de um limiar fixo: se $P(t) \geq 0{,}5$, o sinal é definido como 1 (comprar); caso contrário, é definido como 0 (não operar). Os sinais produzidos nas Q rodadas são concatenados, formando a série temporal completa de sinais de negociação *Trading Signals* (TS) para aquela ação. Por exemplo, para 123 rodadas com 5 dias cada, a série TS contém 615 sinais.

O procedimento completo é formalizado no Algoritmo 2.

Algoritmo 2 Algoritmo de descoberta de *Trading Signal* (TS)

Entrada: Símbolos de ações, d_1 (tamanho do conjunto de treinamento WFA), d_2 (janela deslizante WFA)
Saída: Série temporal de sinais de negociação TS

$N \leftarrow$ Comprimento da lista de símbolos de ações
for $i = 1$ **até** N **do**
  Obtenha o conjunto de dados D_i da i-ésima ação
  $L \leftarrow$ número de linhas de D_i ▷ Número de observações
  $Q \leftarrow$ Eq. (57) ▷ Número de rodadas WFA
  $H \leftarrow$ Eq. (58) ▷ Observações descartadas para alinhamento
  $TS \leftarrow \emptyset$ ▷ Vetor de sinais, preenchido a cada rodada
  Selecione da observação $H + 1$ até a observação L de D_i
  for $j = 1$ **até** Q **do**
    Reinicialize os parâmetros do modelo ▷ Modelo novo a cada rodada
    Defina o conjunto de treinamento: observações $1 + d_2(j-1)$ até $d_1 + d_2(j-1)$
    Defina o conjunto de teste: observações $d_1 + d_2(j-1) + 1$ até $d_1 + d_2 \cdot j$
    Treine o modelo de aprendizado de máquina com o conjunto de treinamento
    Aplique o modelo treinado ao conjunto de teste para obter probabilidades P
    $TS_0(t) \leftarrow 1$ se $P(t) \geq 0{,}5$, senão 0 ▷ Sinal binário por limiar
    $TS \leftarrow$ Concatene TS e TS_0 em ordem ▷ Adiciona os d_2 sinais ao final do vetor
  end for
end for
return TS

4.3 *Backtesting* e Cálculo das Métricas de *Performance*

Os modelos candidatos de aprendizado de máquina geram uma série temporal de sinais binários por meio do método *Walk-Forward Analysis* (WFA), denominada *Trading Signal* (TS). Para cada dia de negociação t, o sinal assume o valor 1 (operar) ou 0 (não operar). Quando $TS_t = 1$, compra-se a ação no fechamento do dia t e vende-se no fechamento do dia $t+1$; quando $TS_t = 0$, nenhuma operação é realizada. Assim, considera-se uma estratégia *long-only* com período de detenção de 1 dia. Assume-se, ainda, a possibilidade de executar as ordens ao preço de fechamento, o que é uma simplificação comumente utilizada em estudos acadêmicos; entretanto, a probabilidade de uma transação ocorrer a um preço muito próximo do fechamento (da ordem de 1 a 2 *ticks*) é alta, conforme discutido em Lv et al. (2023).

A utilização de dados históricos para simular a execução de uma estratégia de negociação é denominada *backtesting*. Na fase de pesquisa e desenvolvimento de modelos de negociação, séries históricas são empregadas para estimar o desempenho teórico de estratégias, permitindo comparar abordagens em condições controladas (NI; ZHANG, 2005).

O ponto de partida do *backtesting* é o cálculo da taxa de retorno diária do ativo. Dado o preço de fechamento C_t no dia t, a taxa de retorno diário TRD_t é definida como a

variação percentual entre dois fechamentos consecutivos:

$$TRD_t = \frac{C_t - C_{t-1}}{C_{t-1}}. \tag{59}$$

Essa grandeza corresponde ao retorno da estratégia *Buy-and-Hold* (comprar e manter) no horizonte de 1 dia e representa o retorno bruto disponível no mercado a cada pregão, independentemente de qualquer regra de negociação.

A partir de TRD_t, calcula-se a taxa de retorno diária do modelo de negociação, $TRMD_t$, que incorpora o sinal produzido pelo modelo candidato. A formulação é:

$$TRMD_t = TS_{t-1} \cdot TRD_t. \tag{60}$$

Observa-se que o sinal utilizado é o do dia anterior (TS_{t-1}) e não o do dia corrente. Isso reflete a dinâmica operacional: ao final do dia $t-1$, o investidor observa o sinal emitido pelo modelo e decide se comprará a ação no fechamento do dia. O retorno capturado corresponde, portanto, à variação do fechamento de $t-1$ para o de t, que é exatamente TRD_t. Quando $TS_{t-1} = 1$, o investidor permanece posicionado e captura integralmente o retorno do dia, seja ele positivo ou negativo. Quando $TS_{t-1} = 0$, o investidor fica fora do mercado e seu retorno é nulo naquele pregão.

A série $TRMD_t$ possibilita calcular oito métricas de desempenho que avaliam a qualidade da estratégia sob diferentes perspectivas: taxa de acerto (do inglês *win rate*, WR), taxa de retorno anualizada (do inglês *annualized return rate*, ARR), índice de Sharpe anualizado (do inglês *annualized Sharpe ratio*, ASR), *drawdown* máximo (do inglês *maximum drawdown*, MDD), índice de Calmar (do inglês *Calmar ratio*, CAR) (CHEN; GUESTRIN, 2016), índice de Sortino (do inglês *Sortino ratio*, SOR), índice de Sharpe modificado (do inglês *modified Sharpe ratio*, MSR) e tempo médio de recuperação (do inglês *average recovery time*, ART). As definições formais são apresentadas a seguir.

A taxa de acerto (WR) é a razão entre o número de pregões com retorno positivo e o número total de pregões considerados no *backtesting*:

$$\text{WR} = \frac{\text{Quantidade de dias com lucro}}{\text{Quantidade total de dias de negociação}}. \tag{61}$$

A WR mede a proporção de negociações lucrativas geradas pela estratégia. Como métrica básica, ela permite analisar a consistência dos sinais emitidos ao longo do período avaliado.

A taxa de retorno anualizada (ARR) é obtida a partir do crescimento do valor líquido (*net value*) ao longo do horizonte de investimento. Suponha que n_t e n_{t+h} representem o

valor líquido nos instantes t e $t+h$, respectivamente, onde h é o período total de investimento e m é o número de períodos por ano. A definição formal é:

$$\text{ARR} = \left(\frac{n_{t+h}}{n_t}\right)^{m/h} - 1. \tag{62}$$

O índice de Sharpe anualizado (ASR) mede o retorno ajustado ao risco. Considere que, no horizonte h, a taxa anualizada seja ARR_h, o desvio padrão das taxas de retorno seja σ_h e r_f seja a taxa livre de risco. A definição formal é:

$$\text{ASR} = \sqrt{m}\,\frac{\text{ARR}_h - r_f}{\sigma_h}. \tag{63}$$

O *drawdown* máximo (MDD) quantifica a maior perda percentual a partir de um pico histórico do valor líquido ao longo do horizonte avaliado. Seja n_t o valor líquido no tempo t. Define-se o *drawdown* em τ como a queda em relação ao maior valor observado antes de τ; então, o MDD é o maior *drawdown* no intervalo:

$$D_\tau = \max\left(0,\, \max_{t\in(0,\tau)} n_t - n_\tau\right); \quad \text{MDD}_h = \max_{\tau\in[0,h]} \frac{D_\tau}{\max_{t\in(0,\tau)} n_t}. \tag{64}$$

O índice de Calmar (CAR) é a razão entre a taxa de retorno anualizada e o *drawdown* máximo:

$$\text{CAR} = \frac{\text{ARR}}{\text{MDD}}. \tag{65}$$

O índice de Sortino (SOR) é uma variação do índice de Sharpe que utiliza apenas a volatilidade dos retornos inferiores a uma taxa mínima aceitável τ. Seja $r = (r_1, r_2, \ldots, r_Z)$ a série de retornos no horizonte considerado e Z é o número total de retornos na série. A definição formal é:

$$\text{SOR} = \frac{E(r) - \tau}{\sqrt{\dfrac{1}{Z}\displaystyle\sum_{t=1}^{Z} \max(0, \tau - r_t)^2}}. \tag{66}$$

O índice de Sharpe modificado (MSR) ajusta o ASR ao incorporar penalizações associadas à assimetria (*skewness*) e à curtose (*kurtosis*) da distribuição dos retornos. Seja S a assimetria e Ks a curtose. A definição formal é:

$$\text{MSR} = \text{ASR}\left(1 + \frac{S}{6}\,\text{ASR} - \frac{Ks - 3}{24}\,\text{ASR}^2\right). \tag{67}$$

O tempo médio de recuperação (ART) mede o tempo necessário para que o valor do investimento retorne ao seu nível de pico após um *drawdown*. O cálculo do ART envolve identificar os k eventos de *drawdown* totalmente recuperados na série. Seja I_t o valor do investimento no tempo t. Define-se o i-ésimo tempo de recuperação T_i como:

$$T_i = \min\{t > t^{(i)}_{\text{trough}} \mid I_t \geq H_{t^{(i)}_{\text{start}}}\} - t^{(i)}_{\text{trough}}, \tag{68}$$

onde H_t é o *High Water Mark* (pico histórico), definido como $\max_{0\leq\tau\leq t} I_\tau$, $t^{(i)}_{\text{start}}$ é o instante de início do i-ésimo *drawdown* e $t^{(i)}_{\text{trough}}$ é o instante do fundo do vale do i-ésimo *drawdown*. Dessa forma, o ART é a média aritmética dos tempos de recuperação:

$$\text{ART} = \frac{1}{k}\sum_{i=1}^{k} T_i. \tag{69}$$

O procedimento de *backtesting* é aplicado a cada uma das N ações da amostra. Para cada ação i, as oito métricas são calculadas a partir da série $TRMD_t$, resultando em um valor escalar por métrica (por exemplo, $\text{WR}[i] = 0{,}53$ indica que a estratégia obteve retorno diário positivo em 53% dos pregões). Esses valores são armazenados em vetores de tamanho N e utilizados na etapa de recomendação TODIM: na abordagem setorial, calcula-se a média por setor; na abordagem individual, utilizam-se os valores diretamente. O procedimento completo é formalizado no Algoritmo 3.

Algoritmo 3 Algoritmo de *backtesting* para estratégia de negociação

Entrada: *Trading Signal* TS, lista de símbolos de ações
Saída: Vetores de métricas de desempenho para as N ações

$N \leftarrow$ Tamanho da lista de símbolos de ações
Inicialize os vetores $WR, ARR, ASR, MDD, CAR, SOR, MSR, ART$ como vazios
for $i = 1$ **até** N **do**
    Obtenha os dados OHLCV da i-ésima ação
    Obtenha o preço de fechamento C_t
    Obtenha o *Trading Signal* TS_t do modelo candidato
    Calcule TRD_t conforme Eq. (59) ▷ Retorno diário do ativo
    Calcule $TRMD_t$ conforme Eq. (60) ▷ Retorno filtrado pelo sinal
    Calcule as 8 métricas sobre a série $TRMD_t$ da ação i
    Armazene $WR[i], ARR[i], ASR[i], MDD[i], CAR[i], SOR[i], MSR[i], ART[i]$ nos respectivos vetores
end for
return Vetores $WR, ARR, ASR, MDD, CAR, SOR, MSR, ART$

Ao final da execução do Algoritmo 3, cada vetor de métrica contém N valores (um para cada ação). Por exemplo, $ASR = (ASR[1], ASR[2], \ldots, ASR[N])$ armazena o índice de Sharpe anualizado de cada ação para um modelo candidato específico. Esse

procedimento é repetido para cada um dos 12 modelos, resultando em 12 conjuntos de vetores de métricas que alimentam a etapa subsequente de recomendação multicritério.

4.4 Recomendação do Modelo Ótimo

Nesta Seção, propõe-se um algoritmo para recomendar o modelo de negociação mais adequado, tanto em granularidade setorial quanto em granularidade individual por ação, com base nas oito métricas de desempenho definidas na Seção 4.3. No método proposto, os indicadores SOR, CAR e MSR são considerados mais próximos das condições reais de negociação e, portanto, recebem maior importância relativa do que WR, ARR, ASR, ART e MDD (LV et al., 2023). Além disso, esses indicadores incorporam medidas associadas a perdas e riscos extremos, o que tende a favorecer modelos mais robustos em cenários reais. Por fim, ressalta-se que a abordagem é descritiva, e não normativa: o objetivo é modelar *como* decisões podem ser tomadas com base em múltiplos critérios, e não prescrever *como* deveriam ser tomadas.

O problema é formulado como um caso de MCDM (*Multi-Criteria Decision Making*): há $M = 8$ critérios (métricas de desempenho) e $N = 12$ alternativas (modelos candidatos). Assim, selecionar o modelo ideal dentre os 12 candidatos, com base nos oito critérios, constitui o núcleo decisório deste trabalho. Vale enfatizar que os 12 modelos candidatos são treinados duas vezes, uma para cada conjunto de dados (com e sem seleção de atributos). No entanto, no problema de decisão multicritério, consideram-se apenas os modelos com melhor desempenho em cada conjunto de dados. Assim, se o modelo LSTM, treinado nos dois conjuntos, apresentar melhor desempenho no conjunto sem seleção de atributos, serão utilizadas as métricas desse LSTM. Por sua vez, se o modelo RNN apresentar melhor desempenho no conjunto com seleção de atributos, serão utilizadas as métricas desse RNN. Dessa forma, o número de modelos analisados é reduzido de 24 para apenas os 12 candidatos descritos. Para resolver esse problema, utiliza-se o método TODIM, que se baseia em comparações par a par entre alternativas, quantificando a vantagem relativa de um modelo em relação a outro em cada critério.

As ações são agrupadas por setor econômico. No caso da B3, adota-se a classificação setorial com $S = 11$ setores: Bens Materiais (BM), Consumo Cíclico (CC), Consumo Não Cíclico (CNC), Comunicações (COM), Financeiro (FIN), Governo, Construção e Serviços (GCS), Óleo, Gás e Biocombustíveis (OGB), Utilidade Pública (UTI), Tecnologia da Informação (TI), Saúde (HLT) e Outros (OTH). Cada setor contém um número variável de ações, totalizando $N = 364$ ações (Tabela 2). Ressalta-se que a Tabela 2 reflete a configuração reportada para 2026 na B3; em outros períodos, a composição setorial e a quantidade de ações podem diferir. A premissa subjacente é que ações do mesmo setor compartilham fatores macroeconômicos e regulatórios, podendo apresentar dinâmicas de

preço mais similares, o que motiva a recomendação de um modelo de *trading* por setor e, adicionalmente, a investigação de um modelo melhor por ação individual.

Tabela 2 – Total de Ações por Setor na B3 em 2026

Setor	**Total de Ações**
BM	31
CC	89
CNC	28
COM	7
FIN	54
GCS	54
OGB	13
UTI	43
TI	15
HLT	22
OTH	8
TOTAL GERAL	**364**

Fonte: (B3, 2026).

A seguir, descreve-se o procedimento TODIM utilizado para construir o ranking dos modelos candidatos.

Considere uma matriz de decisão $D = (d_{ij}) \in \mathbb{R}^{N\times M}$, em que $A = \{A_1, A_2, \ldots, A_N\}$ representa o conjunto de modelos candidatos (alternativas) e $C = \{C_1, C_2, \ldots, C_M\}$ representa o conjunto de critérios (métricas) (Tabela 3). Como os critérios podem ter escalas e direções distintas de preferência (benefício ou custo), torna-se necessário normalizar cada coluna de D.

Tabela 3 – Matriz de decisão multicritério.

	C_1	C_2	$\cdots$	C_M
A_1	d_{11}	d_{12}	$\cdots$	d_{1M}
A_2	d_{21}	d_{22}	$\cdots$	d_{2M}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
A_N	d_{N1}	d_{N2}	$\cdots$	d_{NM}

Passo 1 (Normalização e consistência): realiza-se a normalização por máximo-mínimo em cada critério C. Para critérios de benefício (quanto maior, melhor), utiliza-se:

$$n_{ij} = \frac{d_{ij} - \min_i d_{ij}}{\max_i d_{ij} - \min_i d_{ij}} \quad (i = 1, 2, 3 \ldots, N). \tag{70}$$

Para critérios de custo (quanto menor, melhor), utiliza-se:

$$n_{ij} = \frac{\max_i d_{ij} - d_{ij}}{\max_i d_{ij} - \min_i d_{ij}} \quad (i = 1, 2, 3 \ldots, N). \tag{71}$$

Passo 2 (Pesos dos critérios): seja q_j a importância relativa do critério C_j, para $j = 1, 2, \ldots, M$. Para obter pesos normalizados, aplica-se:

$$w_j = \frac{q_j}{\max(q_1, q_2, \ldots, q_M)}. \tag{72}$$

Passo 3 (Comparação par a par por critério): para quaisquer $i, k \in \{1, 2, \ldots, N\}$ e $j \in \{1, 2, \ldots, M\}$, calcula-se o grau de vantagem de A_i sobre A_k no critério C_j por meio de uma função de valor inspirada na Teoria do Prospecto (TVERSKY; KAHNEMAN, 1992). Adota-se:

$$\Phi_j(A_i, A_k) = \begin{cases} (w_j(n_{ij} - n_{kj}))^{\alpha}, & \text{se } n_{ij} > n_{kj}, \\ 0, & \text{se } n_{ij} = n_{kj}, \\ -\lambda \left(\dfrac{(n_{kj} - n_{ij})}{w_j} \right)^{\beta}, & \text{se } n_{ij} < n_{kj}, \end{cases} \tag{73}$$

em que $0 < \alpha < 1$ e $0 < \beta < 1$ controlam a concavidade/convexidade da função e $\lambda > 1$ representa aversão a perdas. Neste trabalho, utiliza-se $\alpha = \beta = 0.88$ e $\lambda = 2.25$, conforme (TVERSKY; KAHNEMAN, 1992).

Passo 4 (Vantagem total A_i vs. A_k): soma-se a vantagem por critério:

$$\Phi(A_i, A_k) = \sum_{j=1}^{M} \Phi_j(A_i, A_k). \tag{74}$$

Passo 5 (Desempenho global por alternativa): agrega-se a vantagem de A_i em relação a todas as alternativas:

$$\Phi(A_i) = \sum_{k=1}^{N} \Phi(A_i, A_k). \tag{75}$$

Passo 6 (Normalização do escore global): aplica-se máximo–mínimo ao escore global, gerando $\xi(A_i) \in [0,1]$:

$$\xi(A_i) = \frac{\Phi(A_i) - \min_k \Phi(A_k)}{\max_k \Phi(A_k) - \min_k \Phi(A_k)}. \tag{76}$$

Passo 7 (Ranking): os modelos candidatos são ordenados em ordem decrescente de $\xi(A_i)$, isto é, $A_i \succ A_k \iff \xi(A_i) > \xi(A_k)$. Pode-se ler como A_i domina A_k.

Dessa forma, propõe-se um algoritmo unificado para selecionar o modelo de negociação ideal, aplicável tanto à granularidade setorial quanto à granularidade individual por ação, conforme o Algoritmo 4. Quando $\mathcal{U}$ corresponde ao conjunto de setores, constrói-se

D_u pela média aritmética das métricas das ações dos setores; quando $\mathcal{U}$ corresponde ao conjunto de ações individuais, D_u contém diretamente as métricas dessas ações.

Algoritmo 4 Algoritmo de recomendação TODIM para modelo de negociação ideal

Entrada: Importâncias $q = (q_1, \ldots, q_M)$; N modelos candidatos; M critérios; lista de unidades $\mathcal{U}$ (setores ou ações individuais)
Saída: Ranking dos N modelos para cada unidade $u \in \mathcal{U}$

Melhor-A $\leftarrow$ lista vazia ▷ Armazena o ranking por unidade
for all $u \in \mathcal{U}$ **do**
    Construa $D_u = (d_{ij}) \in \mathbb{R}^{N \times M}$ ▷ Média setorial ou métricas diretas da ação
    Normalize D_u obtendo $N_u = (n_{ij}) \in \mathbb{R}^{N \times M}$ conforme Eq. (70) e Eq. (71)
    Calcule w_j conforme Eq. (72)
    for $i = 1$ **até** N **do**
        $\Phi(A_i) \leftarrow 0$
        for $k = 1$ **até** N **do**
            Calcule $\Phi(A_i, A_k)$ conforme Eq. (74)
            $\Phi(A_i) \leftarrow \Phi(A_i) + \Phi(A_i, A_k)$
        end for
    end for
    Normalize os escores $\Phi(A_i)$ obtendo $\xi(A_i)$ conforme Eq. (76)
    Ordene os modelos por $\xi(A_i)$ em ordem decrescente
    Armazene o ranking ordenado em *Melhor-A*$[u]$
end for
return *Melhor-A*

Sendo assim, a abordagem proposta neste trabalho investiga duas granularidades de recomendação:

- **Melhor modelo por setor:** as métricas são agregadas pela média aritmética das ações de cada setor, formando uma matriz de decisão (12 modelos $\times$ 8 métricas) por setor. O TODIM é aplicado a essa matriz para gerar um ranking e identificar o modelo ótimo em cada setor.
- **Melhor modelo por ação individual:** as métricas de cada ação são utilizadas diretamente, sem agregação setorial. O TODIM é aplicado à matriz de decisão de cada ação, identificando o modelo ótimo para essa ação.

A comparação entre as duas granularidades permite testar a premissa central do trabalho de referência: se a classificação setorial é uma simplificação suficiente ou se a heterogeneidade intrassetorial justifica uma recomendação mais granular. Por fim, testa-se essa hipótese comparando o desempenho (retorno e demais métricas) entre: (i) o uso do melhor modelo por ação, (ii) o melhor modelo por setor, (iii) o índice de mercado, (iv) a estratégia *buy-and-hold*, em um período predefinido e (v) estratégias de portfólio *long-short* baseadas nas recomendações setoriais e individuais dos modelos.

5 Experimentos e Resultados

Neste capítulo, apresentam-se, de forma estruturada, o ambiente experimental, as análises empíricas e a validação do sistema de recomendação desenvolvido neste trabalho. Inicialmente, detalha-se a configuração experimental, o que compreende a descrição do conjunto de dados B3ICS, a aplicação da metodologia *Walk-Forward Analysis* (WFA) e a divisão espaço-temporal que estrutura a avaliação em quatro cenários distintos. Adicionalmente, expõe-se o fluxo de seleção de atributos baseado no filtro *Pearson Redundancy Based Filter* - PRBF (que combina correlação de Pearson e informação mútua), a parametrização dos 12 modelos candidatos de aprendizado de máquina e de aprendizado profundo, bem como as métricas matemáticas utilizadas para aferir o desempenho de negociação e a capacidade de generalização algorítmica.

Na sequência, o capítulo aborda os resultados experimentais divididos em quatro eixos fundamentais. O primeiro avalia, por meio do teste de Wilcoxon pareado, o impacto da redução dimensional no desempenho das arquiteturas. O segundo eixo analisa o desempenho de negociação dos modelos. O terceiro eixo detalha a aplicação do algoritmo de decisão multicritério TODIM para identificar o modelo de negociação ótimo nas granularidades setorial e individual, mensurando a capacidade de generalização do sistema por meio de métricas de recuperação e de ordenação. Por fim, investiga-se o valor prático do ecossistema de recomendação por meio da construção e avaliação de portfólios *long-short* gerados a partir de diferentes esquemas de alocação de capital, evidenciando a eficiência do método na mitigação de riscos sistêmicos.

5.1 Configuração Experimental

5.1.1 Descrição do Conjunto de Dados

Nesta Seção, adota-se o B3ICS (B3 *Index Constituent Stocks*) como conjunto de dados principal, composto por 239 ações distribuídas em 8 setores econômicos da B3 (Tabela 5). Essa composição corresponde ao subconjunto efetivamente disponível no período experimental analisado. A redução em relação ao universo inicial de ativos decorre da disponibilidade efetiva de séries históricas no período considerado, bem como da necessidade de agregar setores com quantidade insuficiente de ações para permitir uma avaliação estatisticamente consistente.

O conjunto de dados brutos (OHLCV) de cada ação abrange o período de janeiro de 2016 a junho de 2020. Contudo, nem toda a extensão temporal é utilizada diretamente pelos modelos de aprendizado de máquina. O ano de 2016 serve exclusivamente como período

para o cálculo dos indicadores técnicos, uma vez que diversos indicadores necessitam de um histórico mínimo de observações para serem computados (por exemplo, o Mass Index requer 25 dias e a EMA de 26 períodos do MACD demanda ao menos 26 observações prévias). Dessa forma, os valores dos indicadores técnicos tornam-se estáveis a partir do início de 2017.

A metodologia WFA utiliza uma janela de treinamento de $d_1 = 250$ dias (aproximadamente um ano). Assim, a primeira rodada do WFA treina o modelo com os 250 dias úteis de 2017 e gera previsões para os 5 dias seguintes, que correspondem ao início de janeiro de 2018. Consequentemente, as séries de sinais de negociação (TS) produzidas pelos modelos cobrem apenas o período de janeiro de 2018 a junho de 2020, totalizando aproximadamente 615 sinais por ação.

É importante ressaltar que o WFA é executado de forma independente para cada uma das 239 ações, sem distinção entre as amostras nesta etapa. Todos os 12 modelos candidatos são treinados e geram sinais para todas as ações indistintamente. A separação entre conjuntos ocorre apenas na etapa posterior de avaliação, quando as ações são divididas em dois grupos para investigar a capacidade de generalização do algoritmo de recomendação:

- **In-Sample (InS):** 190 ações (80% de cada setor), utilizadas para calibrar o algoritmo de recomendação TODIM nº1;
- **Out-of-Sample (OutS):** 49 ações (20% de cada setor), reservadas para verificar se o modelo recomendado com base nas ações InS generaliza para ações não vistas durante a calibração.

Adicionalmente, o eixo temporal é dividido em dois períodos para avaliar a estabilidade das recomendações ao longo do tempo:

- **Treinamento (TraD):** de janeiro de 2018 a dezembro de 2019, período utilizado para calcular as métricas de desempenho que alimentam o TODIM nº1;
- **Teste (TesD):** de janeiro a junho de 2020, período utilizado para avaliar se o modelo treinado mantém bom desempenho no futuro.

O cruzamento dessas duas dimensões (amostra e tempo) resulta em quatro cenários de avaliação, conforme a Tabela 4:

Tabela 4 – Quatro cenários de avaliação para generalização do algoritmo de recomendação.

Cenário	**Ações**	**Período**	**Finalidade**
InS-TraD	In (190)	2018–2019	Calibração do TODIM nº1
InS-TesD	In (190)	2020	Generalização temporal: TODIM nº2
OutS-TraD	Out (49)	2018–2019	Generalização amostral: TODIM nº3
OutS-TesD	Out (49)	2020	Generalização amostral e temporal: TODIM nº4

O cenário InS-TraD é o cenário principal[1], pois é nele que o TODIM identifica o modelo ótimo para cada setor (ou ação). Os demais cenários testam se essa recomendação se sustenta quando aplicada a ações diferentes (OutS), a um período futuro (TesD) ou a ambos simultaneamente (OutS-TesD)[2].

Tabela 5 – Distribuição de Ações por Setor (B3ICS)

Setor	**In**	**Out**	**Total**
BM	25	6	31
CC	50	13	63
CNC	10	3	13
COM	5	1	6
FIN	32	8	40
GCS	30	8	38
OGB	7	2	9
UTI	31	8	39
TOTAL	**190**	**49**	**239**

5.1.2 Seleção de Atributos: Correlação de Pearson mais Informação Mútua

A estratégia de seleção de atributos adotada neste trabalho fundamenta-se no paradigma de relevância e redundância, princípio consolidado na literatura de seleção de variáveis e formalizado de maneira canônica pelo método *minimum Redundancy Maximum Relevance* (mRMR) proposto por Peng, Long e Ding (2005). Sob esse paradigma, um subconjunto ótimo de atributos é aquele que maximiza a relevância de cada variável em relação à classe alvo e, simultaneamente, minimiza a redundância entre as variáveis selecionadas. A combinação específica empregada nesta dissertação, na qual o coeficiente de correlação de Pearson quantifica a redundância entre atributos e a informação mútua quantifica a relevância de cada atributo em relação ao alvo, encontra respaldo direto na literatura, uma vez que essa associação de métricas é explicitamente prevista para variáveis contínuas no âmbito do paradigma mRMR Radovic et al. (2017).

Cabe esclarecer que a seleção de atributos não constitui o objeto central desta pesquisa, mas sim uma etapa instrumental de um arcabouço mais amplo de recomendação de modelos de negociação. O subconjunto de atributos selecionado alimenta um conjunto de algoritmos de aprendizado de máquina, treinados para classificar a direção do preço de cada ativo. Nesse fluxo, a seleção de atributos tem a função de fornecer um espaço de variáveis comum, conciso e não redundante, que serve de insumo padronizado para

[1] Para melhor entendimento, vale dizer que existem quatro execuções do TODIM: a primeira referente ao cenário InS-TraD, a segunda ao cenário InS-TesD, a terceira ao OutS-TraD e a quarta ao OutS-TesD. A ideia é verificar se o TODIM n^o 1 (InS-TraD) consegue generalizar bem aos demais cenários.

[2] A divisão das ações em 80% In e 20% Out foi realizada de forma aleatória em cada setor, com semente fixa para reprodutibilidade.

todos os modelos, assegurando a comparabilidade entre eles. Por essa razão, privilegiou-se um procedimento de filtro consagrado e de baixo custo computacional, cuja eventual sofisticação adicional não se justificaria diante do papel auxiliar que a seleção desempenha no escopo geral do trabalho.

A operacionalização desse paradigma foi conduzida por meio de um procedimento sequencial de duas etapas, em conformidade com o desenho do filtro *Pearson Redundancy Based Filter* (PRBF) de Biesiada e Duch (2007), que estrutura a seleção em uma fase de tratamento da redundância e outra de avaliação da relevância. Na primeira etapa, calcula-se a matriz de correlação de Pearson entre todos os pares de atributos e, para cada par cujo valor absoluto do coeficiente excede um limiar predefinido, descarta-se o atributo de menor informação mútua com a classe-alvo, preservando-se, assim, a variável mais informativa de cada par redundante. Na segunda etapa, os atributos remanescentes são ordenados em ordem decrescente de informação mútua com o alvo, sendo retidos os de maior relevância. Cabe ressaltar que o trabalho recente de Gong et al. (2024), ao propor a integração de uma medida linear (correlação de Pearson) e de uma medida não linear (informação mútua) em uma única função de avaliação, reforça a pertinência de combinar indicadores de naturezas distintas para uma caracterização mais abrangente das dependências entre variáveis.

Ressalta se que tanto a matriz de correlação quanto os *escores* de informação mútua foram calculados sobre uma base de dados única, obtida pelo empilhamento (*pooling*) das observações diárias de todas as N ações da amostra, e não por meio da média de valores calculados individualmente para cada ativo. Nessa base agregada, cada linha corresponde à observação de um indicador técnico em um dia de negociação de uma ação específica, associada ao respectivo rótulo binário de direção do preço no dia seguinte, de modo que cada atributo recebeu um único valor de correlação e de informação mútua, representativo do conjunto completo de ações.

O limiar de correlação foi fixado em $|\rho| > 0{,}85$ para identificar pares redundantes. Esse valor situa-se na faixa de correlação classificada como muito forte pelas escalas usualmente adotadas na literatura para interpretação da magnitude do coeficiente de correlação, que consideram associações com valor absoluto a partir de $0{,}8$ como muito fortes ou extremamente fortes Hinkle, Wiersma e Jurs (2003), Mukaka (2012). A escolha de um patamar ligeiramente superior a esse piso configura um critério deliberadamente conservador: ao restringir a remoção aos pares de atributos cuja associação linear é muito forte, ou seja, quase colineares, minimiza-se o risco de descartar variáveis que carregam informação apenas parcialmente sobreposta. Convém observar que tais escalas foram concebidas para a interpretação da força da associação, e não como critério formal de redundância; sua adoção aqui serve, portanto, de referência para fundamentar a ordem de grandeza do limiar empregado, cuja calibração final é corroborada pela estabilidade do subconjunto selecionado discutida adiante. Para a etapa de relevância, adotou-se um

critério de corte relativo, no qual foram descartados os atributos cuja informação mútua com o alvo era inferior a 10% do maior valor observado, eliminando-se a cauda de variáveis com contribuição informacional desprezível. A opção por um patamar relativo, em detrimento de um valor absoluto, torna o critério adaptável à magnitude do sinal presente nos dados, característica especialmente relevante em contextos de previsão financeira, nos quais a informação mútua entre atributos e alvo tende a assumir valores reduzidos (GUNDUZ; CATALTEPE, 2015).

A adequação do corte adotado pode ser verificada graficamente na Figura 5, que apresenta a informação mútua das variáveis, ordenada decrescentemente. Observa-se que a curva apresenta um decaimento acentuado nas primeiras posições, seguido de uma região de inflexão, a partir da qual os valores passam a decrescer gradualmente até zero. O limiar de relevância coincide aproximadamente com essa região de inflexão, de modo que o ponto de corte adotado não decorre exclusivamente do critério relativo a 10%, mas também é corroborado pela inspeção visual da curva de relevância. Convém ainda observar que, dada a proximidade dos valores de informação mútua na vizinhança do corte, pequenas variações no número de atributos retidos não alteram substancialmente o conjunto selecionado, o que confere robustez à escolha realizada.

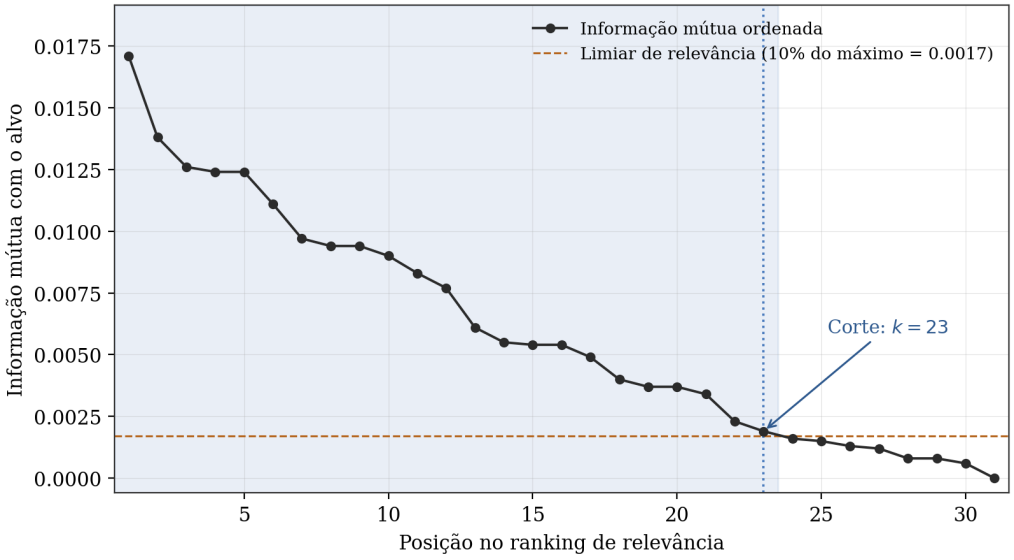

Figura 5 – Informação mútua de cada atributo em relação à variável-alvo, ordenada decrescentemente. A linha tracejada indica o limiar de relevância correspondente a 10% do valor máximo observado; a região sombreada destaca os $k = 23$ atributos retidos. O ponto de corte coincide com a região de inflexão da curva, evidenciando a separação entre os atributos informativos e a cauda de contribuição desprezível.

Cabe justificar a ausência de tratamentos tradicionais de séries temporais, como testes de estacionariedade, diferenciação ou decomposição de tendência, no caso dos indicadores técnicos utilizados como variáveis de entrada. Em primeiro lugar, o problema

abordado neste trabalho é de *classificação binária* (prever se o ativo subirá ou descerá no dia seguinte), e não de previsão de séries temporais univariadas, contexto para o qual metodologias como Box-Jenkins (ARIMA/SARIMA) foram concebidas (BOX; JENKINS; REINSEL, 2013). Os indicadores técnicos não constituem a variável a ser prevista, mas sim atributos de entrada de um classificador; portanto, o requisito relevante é que seus valores sejam informativos para a separação das classes, e não que obedeçam a pressupostos de estacionariedade estrita. Em segundo lugar, a grande maioria dos 44 indicadores já é estacionária por construção: osciladores como RSI, Williams %R e Estocástico K/D/J possuem domínio limitado (e.g., $[0, 100]$ ou $[-100, 0]$); indicadores baseados em retorno (ROC, Ret, Return) tendem a ser estacionários por definição; e indicadores de volatilidade (ATR, Sigma, VOLA) oscilam em torno de um nível médio. Os poucos indicadores potencialmente não estacionários (OBV, AD, SMA, EMA) são mapeados para o intervalo $[0, 1]$ pela normalização min-max aplicada na etapa de construção do *dataset*.

5.1.3 Resultados da Seleção de Atributos

A seleção de atributos foi aplicada ao conjunto agregado de todas as 239 ações, totalizando aproximadamente 197.500 observações diárias. O conjunto original de 44 indicadores técnicos (normalizados por min-max) passou pelos dois estágios descritos na Seção 4.1, cujos resultados são detalhados a seguir.

Estágio 1: Remoção de redundância por Correlação de Pearson

O primeiro estágio identificou 15 pares de indicadores com correlação absoluta superior ao limiar $|\rho| > 0{,}85$. Para cada par, o indicador com menor Informação Mútua para o rótulo binário foi removido, preservando o indicador mais informativo. A Tabela 6 apresenta os pares identificados, seus coeficientes de correlação e o indicador removido em cada caso.

Os resultados revelam padrões esperados de redundância entre os indicadores. O grupo estocástico (WR, RSV, Kvalue, Jvalue) apresenta correlações superiores a 0,89, uma vez que todos derivam da posição relativa do preço de fechamento no intervalo recente. As médias móveis (SMA_5, SMA_10, EMA_5, EMA_10) apresentam correlações acima de 0,99, pois diferem apenas no período e no esquema de ponderação. O par SONAR/SONSIG ($\rho = 0{,}981$) reflete o fato de que o SONSIG é simplesmente uma EMA do SONAR. Ao final deste estágio, 13 indicadores foram removidos e 31 permaneceram.

Estágio 2: Seleção por Informação Mútua

No segundo estágio, a Informação Mútua (MI) entre cada um dos 31 indicadores sobreviventes e o rótulo binário foi calculada. O limiar automático de 10% da MI máxima resultou na seleção de 23 indicadores. A informação mútua foi estimada por meio do algoritmo baseado em k-vizinhos mais próximos (k = 5), tratando os atributos como variáveis

Tabela 6 – Pares de indicadores com $|\rho| > 0{,}85$ e indicador removido por menor MI.

Ind. 1	**Ind. 2**	$\lvert\rho\rvert$	**MI Ind. 1**	**MI Ind. 2**	**Removido**
WR	RSV	0,989	0,0098	0,0094	RSV
WR	Kvalue	0,890	0,0098	0,0011	Kvalue
WR	Jvalue	0,949	0,0098	0,0002	Jvalue
BandPer	CCI	0,877	0,0020	0,0038	BandPer
DIS	MOM	0,940	0,0018	0,0084	DIS
DIS	ROC	0,948	0,0018	0,0058	DIS
DIS	Return	0,922	0,0018	0,0062	DIS
MAO	TRIX	0,850	0,0015	0,0005	TRIX
MAO	MACD	0,936	0,0015	0,0008	MACD
MOM	ROC	0,925	0,0084	0,0058	ROC
MOM	Return	0,862	0,0084	0,0062	Return
SONAR	SONSIG	0,981	0,0019	0,0011	SONSIG
SMA_5	SMA_10	0,995	0,0055	0,0041	SMA_10
SMA_5	EMA_5	0,999	0,0055	0,0033	EMA_5
SMA_5	EMA_10	0,996	0,0055	0,0043	EMA_10

contínuas, com semente fixa para reprodutibilidade. A Tabela 7 apresenta o ranking completo, incluindo os 23 indicadores selecionados e os 8 descartados.

Tabela 7 – Ranking dos indicadores por Informação Mútua. Os 23 indicadores acima da linha tracejada foram selecionados (MI $\geq$ 0,0017).

Rank	**Indicador**	**MI**	**Rank**	**Indicador**	**MI**
1	OBV	0,0171	13	RMI	0,0061
2	AD	0,0138	14	CMF	0,0055
3	EOM	0,0126	15	NBIAS	0,0054
4	FI	0,0124	16	SMA_5	0,0054
5	RSI	0,0124	17	ATR	0,0049
6	NCO	0,0111	18	MFI	0,0040
7	VMA	0,0097	19	CCI	0,0037
8	WR	0,0094	20	MI	0,0037
9	VOLA	0,0094	21	Sigma	0,0034
10	Ret	0,0090	22	CO	0,0023
11	MOM	0,0083	23	SONAR	0,0019
12	PSY	0,0077			
Indicadores descartados (MI $<$ 0,0017)					
24	ADX	0,0016	28	BandWid	0,0008
25	MAO	0,0015	29	Dvalue	0,0008
26	SROC	0,0013	30	PO	0,0006
27	VROC	0,0012	31	VOS	0,0000

Os indicadores de volume (OBV, AD, FI, VMA, CMF) e os indicadores de momentum (RSI, NCO, MOM, Ret) dominam as primeiras posições do ranking, sugerindo que a informação sobre fluxo de volume e velocidade das variações de preço é mais relevante para a classificação binária do que os indicadores de tendência suavizada. O VOS (Volume

Oscillator) apresentou MI igual a zero, indicando ausência de dependência estatística em relação ao rótulo.

A Tabela 8 resume o processo completo de redução dimensional.

Tabela 8 – Resumo do processo de seleção de atributos.

Etapa	**Indicadores**
Conjunto original	44
Removidos por correlação ($\|\rho\| > 0{,}85$)	-13
Sobreviventes pós-correlação	31
Removidos por MI ($< 10\%$ da MI máx.)	-8
Conjunto final selecionado	**23**

O processo de seleção reduziu a dimensionalidade em aproximadamente 48%, de 44 para 23 indicadores. Os *datasets* filtrados foram gerados para cada uma das 239 ações, contendo a data, as 23 *features* selecionadas e o rótulo binário, totalizando 25 colunas. Esses *datasets* são utilizados para treinar os modelos candidatos na configuração com seleção de atributos, cujos resultados são comparados aos da configuração completa de 44 indicadores, apresentados na Seção 5.2.1.

5.1.4 Configuração dos Modelos

Os 12 modelos candidatos descritos na Seção 4.2 foram configurados com base nos parâmetros adotados por Lv et al. (2023). Os atributos de entrada para os modelos tradicionais são organizados como matrizes bidimensionais de dimensão $(250, p)$, onde p é o número de *features* (44 no *dataset* completo ou 23 no *dataset* com seleção de atributos). Para os modelos de aprendizado profundo com arquitetura recorrente (RNN, LSTM e GRU), os dados são estruturados como tensores tridimensionais de dimensão $(1, 250, p)$, onde a primeira dimensão representa o *batch* e a segunda os passos de tempo. Os modelos densos (MLP, DBN e SAE) utilizam o mesmo formato matricial que os modelos tradicionais. Ademais, vale mencionar que os modelos foram implementados na linguagem *Python*[3].

Ainda mais, a mitigação do sobreajuste (*overfitting*) constitui um desafio central na modelagem preditiva de séries temporais financeiras, sendo estruturada neste trabalho em três níveis metodológicos complementares: o tratamento dimensional, a ativação de mecanismos de regularização intrínsecos às arquiteturas dos modelos e a adoção de um protocolo de validação cronológica estrita. Inicialmente, a redução da complexidade do espaço de entrada, por meio da seleção de atributos com base na Informação Mútua, atuou como um mecanismo primário de regularização. Ao reduzir a dimensionalidade dos dados, esse filtro restringiu o fluxo de informações aos sinais de maior relevância, protegendo

[3] Todos os códigos utilizados e os dados obtidos podem ser consultados em <https://www.kaggle.com/datasets/rodrigomaicon/conjunto-de-dados-b3ics-e-b3ics-mi> e <https://github.com/rodrigopiece/TODIM_TRADING_B3_MODELOS_ML>

algoritmos desprovidos de seleção implícita nativa, como a Regressão Logística, o *Naive Bayes*, o *Multilayer Perceptron* e a *Recurrent Neural Network*, da memorização de relações espúrias e redundantes.

No âmbito algorítmico, os modelos mais complexos valeram-se de propriedades arquiteturais específicas para inibir a memorização do ruído inerente ao mercado. O *Random Forest* e o *XGBoost* utilizaram, respectivamente, o conceito de *bagging* com seleção aleatória de atributos e o princípio de impulsionamento regularizado com taxa de aprendizado restritiva, limitando a dependência excessiva do ruído de treino. Paralelamente, o *Support Vector Machine* demonstrou capacidade de generalização ao maximizar a margem do hiperplano, com flexibilidade para violações moderadas, enquanto arquiteturas baseadas em *autoencoders* forçaram a compressão dos dados para preservar apenas características latentes expressivas. Já nos modelos de processamento sequencial, como *Long Short-Term Memory* e *Gated Recurrent Unit*, o controle do sobreajuste temporal foi viabilizado pelo mecanismo de portas (*gates*), cuja capacidade dinâmica de esquecer oscilações transientes impede que as redes memorizem flutuações curtas locais ao longo da janela de processamento.

A validação metodológica dessa estrutura de defesa foi assegurada pela adoção da *Walk-Forward Analysis* e pelo uso de métricas de generalização rigorosas aplicadas ao método TODIM. A reestimação contínua dos modelos em janelas deslizantes e a avaliação cega e estritamente fora da amostra garantiram que os resultados refletissem a real capacidade preditiva dos algoritmos frente a dados não observados.

Por fim, a Tabela 9 apresenta as configurações dos modelos tradicionais e a Tabela 10 detalha os parâmetros dos modelos de aprendizado profundo.

Tabela 9 – Configurações dos principais parâmetros dos modelos tradicionais de *Machine Learning* (ML).

Modelo	**Principais Parâmetros**
LR	Função de ligação *logit*; sem regularização.
SVM	Kernel RBF (*Radial Basis Function*); custo de violação $C = 10$.
NB	Distribuição gaussiana; probabilidade *a priori* proporcional às classes.
CART	Profundidade máxima da árvore: 10; critério de divisão: coeficiente de Gini.
RF	Número de árvores: 500; atributos por divisão: $\lfloor \sqrt{p} \rfloor$.
XGB	Profundidade máxima: 10; iterações: 15; taxa de aprendizado: 0,3.

5.1.5 Métricas de Avaliação e Generalização

Conforme detalhado na Seção 4.3, utilizam-se as métricas WR, ARR, ASR, MDD, CAR, SOR, MSR e ART para avaliar a capacidade de negociação de um algoritmo de

Tabela 10 – Configurações dos principais parâmetros dos modelos de aprendizado profundo.

Modelo	Camadas Ocultas	Ativação	Taxa Ap.	Batch	Épocas
MLP	[25, 15, 10, 5]	Sigmóide	0,01	200	5
DBN	[25, 15, 10, 5]	Sigmóide	0,01	200	5
SAE	[20, 10, 5]	Sigmóide	0,01	200	5
RNN	[20, 10, 5]	Sigmóide	0,01	200	5
LSTM	[20, 10, 5]	Tanh	0,01	200	5
GRU	[20, 10, 5]	Tanh	0,01	10	1

Nota: *Taxa Ap.* = taxa de aprendizado; *Batch* = tamanho do lote.

aprendizado de máquina. Ao avaliar o desempenho de um algoritmo em um setor específico, calcula-se, primeiramente, o desempenho de negociação do algoritmo para cada ação, individualmente, e, em seguida, o desempenho médio de todas as ações desse setor. Os resultados de desempenho de negociação são apresentados na Seção 5.2.2 e os de generalização na Seção 5.2.6.

As métricas de generalização utilizadas neste trabalho são adaptadas da área de Recuperação de Informação (do inglês *Information Retrieval*, IR), campo da ciência da computação dedicado ao desenvolvimento de sistemas capazes de localizar, ordenar e apresentar documentos ou itens relevantes em resposta a uma consulta do usuário, cujo exemplo mais conhecido são os mecanismos de busca na internet (MANNING; RAGHAVAN; SCHÜTZE, 2008). Nesse contexto original, métricas como Precisão (*Precision*), Revocação (*Recall*), MAP e NDCG foram concebidas para avaliar a qualidade de uma lista ordenada de resultados retornada por um sistema de busca: quão relevantes são os itens recuperados e quão bem posicionados estão os mais relevantes no topo da lista. A transposição dessas métricas para o contexto de recomendação de modelos de *trading*, realizada originalmente por Lv et al. (2023) e adotada neste trabalho, substitui os documentos recuperados pelos modelos de aprendizado de máquina rankeados pelo TODIM, e a consulta do usuário pelo cenário de calibração (InS-TraD): em vez de perguntar “quão bons são os documentos retornados pelo buscador?”, pergunta-se “quão consistente é o ranking de modelos produzido na calibração quando aplicado a cenários de validação distintos?”. Essa analogia é natural, pois em ambos os casos o objeto central de avaliação é uma lista ordenada, e o que se deseja medir é a qualidade dessa ordenação frente a um referencial considerado correto. Ressalta-se que as formulações de MAP e NDCG empregadas neste trabalho seguem a adaptação proposta por Lv et al. (2023) para o problema de recomendação de modelos de negociação, diferindo das definições clássicas originalmente utilizadas na Recuperação de Informação.

Para avaliar a capacidade de generalização do algoritmo de recomendação, projetam-

se métricas que quantificam a consistência entre os rankings gerados no cenário de calibração (InS-TraD) e os gerados nos demais cenários (InS-TesD, OutS-TraD e OutS-TesD). Assume-se que todas as ações estão divididas em $N = 8$ setores e que há $K = 12$ modelos candidatos em cada setor. Vale dizer que as métricas de generalização são aplicadas apenas para a granularidade setorial e não por ação individual.

Para os dados de calibração (InS-TraD), R_n^k representa o conjunto dos k melhores modelos recomendados ($1 \leq k \leq K$) para o n-ésimo setor ($1 \leq n \leq N$). Para os cenários de validação (InS-TesD, OutS-TraD e OutS-TesD), T_n^k representa o conjunto dos k melhores modelos recomendados no n-ésimo setor. As métricas descritas a seguir avaliam a diferença entre R_n^k e T_n^k, verificando se os modelos recomendados na calibração permanecem adequados nos demais cenários.

1. **PR@k** (Taxa de Precisão, do inglês *Precision Rate*): representa a proporção dos modelos recomendados no cenário de calibração que também aparecem entre os top-k do cenário de validação. Na fórmula, $|\cdot|$ denota a cardinalidade do conjunto:

$$\text{PR@}k = \frac{\left|\left(\bigcup_{n=1}^{N} T_n^k\right) \cap \left(\bigcup_{n=1}^{N} R_n^k\right)\right|}{\left|\bigcup_{n=1}^{N} R_n^k\right|}. \tag{77}$$

   Por exemplo, assumindo $N = 2$ e $k = 3$, com $R_1^3 = \{\text{MLP, RNN, RF}\}$, $R_2^3 = \{\text{DBN, RNN, NB}\}$, $T_1^3 = \{\text{MLP, SVM, RF}\}$ e $T_2^3 = \{\text{LSTM, RNN, LR}\}$, a união dos conjuntos de calibração contém 5 modelos distintos, dos quais 3 coincidem com a validação, resultando em PR@3 $= 3/5 = 0{,}60$;

2. **RR@k** (Taxa de Revocação, do inglês *Recall Rate*): representa a proporção dos modelos recomendados no cenário de validação que também aparecem entre os top-k do cenário de calibração:

$$\text{RR@}k = \frac{\left|\left(\bigcup_{n=1}^{N} T_n^k\right) \cap \left(\bigcup_{n=1}^{N} R_n^k\right)\right|}{\left|\bigcup_{n=1}^{N} T_n^k\right|}. \tag{78}$$

   Utilizando o mesmo exemplo anterior, a união dos conjuntos de validação contém 6 modelos distintos, dos quais 3 coincidem com a calibração, resultando em RR@3 $= 3/6 = 0{,}50$;

3. **F1@k**: média harmônica entre PR e RR, que descreve o desempenho geral do algoritmo de recomendação equilibrando precisão e revocação:

$$\text{F1@}k = \frac{2 \cdot \text{PR@}k \cdot \text{RR@}k}{\text{PR@}k + \text{RR@}k}. \tag{79}$$

   Continuando o exemplo, F1@3 $= 2 \times 0{,}60 \times 0{,}50 \,/\, (0{,}60 + 0{,}50) = 0{,}55$;

4. **CR@k** (Taxa de Cobertura, do inglês *Coverage Rate*): descreve a diversidade do algoritmo de recomendação, medindo a proporção de modelos distintos que aparecem entre os top-k recomendados em relação ao total de modelos candidatos:

$$\text{CR@}k = \frac{\left|\bigcup_{n=1}^{N} T_n^k\right|}{K}. \tag{80}$$

Por exemplo, com $T_1^3 = \{\text{MLP, SVM, RF}\}$ e $T_2^3 = \{\text{LSTM, RNN, LR}\}$, a união contém 6 modelos distintos de um total de $K = 12$, resultando em CR@3 $= 6/12 = 0{,}50$;

5. **MRR** (Classificação Recíproca Média, do inglês *Mean Reciprocal Rank*): avalia em que posição do ranking de calibração aparece o melhor modelo recomendado no cenário de validação. Na fórmula, rank_{RT}^n representa a posição do melhor modelo de validação no ranking de calibração do n-ésimo setor:

$$\text{MRR} = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{\text{rank}_{RT}^n}. \tag{81}$$

Por exemplo, com $N = 2$, se o ranking completo de calibração for:

$$R_1^{12} = \{\text{MLP}, \text{RNN}, \text{RF}, \text{NB}, \text{DBN}, \text{SVM}, \text{LR}, \text{SAE}, \text{CART}, \text{LSTM}, \text{GRU}, \text{XGB}\};$$
$$R_2^{12} = \{\text{DBN}, \text{RNN}, \text{NB}, \text{CART}, \text{LSTM}, \text{GRU}, \text{XGB}, \text{MLP}, \text{RF}, \text{SVM}, \text{LR}, \text{SAE}\},$$

e os melhores modelos de validação forem $T_1^1 = \{\text{MLP}\}$ e $T_2^1 = \{\text{LSTM}\}$, então $\text{rank}_{RT}^1 = 1$ e $\text{rank}_{RT}^2 = 5$, resultando em MRR $= (1/1 + 1/5)/2 = 0{,}60$;

6. **MAP** (Média da Precisão Média, do inglês *Mean Average Precision*): é uma variante da revocação sensível à ordem dos resultados. Posições mais altas dos modelos corretos no ranking resultam em valores maiores de MAP:

$$\text{MAP} = \frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K}\frac{\sum_{i=1}^{k} I\{\text{rank}_{RT}[1:k] < k\} + 1}{k}, \tag{82}$$

onde $\text{rank}_{RT}[1:k]$ é o vetor das posições ocupadas, no ranking de calibração, pelos top-k modelos do ranking de validação; e $I\{a < b\}$ é a função indicadora que retorna 1 se $a < b$ e 0 caso contrário. Para cada valor de k, o numerador conta quantos modelos entre os top-k da validação estavam posicionados *acima* da posição k no ranking de calibração (acrescido de 1 para evitar valor nulo quando nenhum modelo satisfaz a condição), e divide pelo total de k. O resultado é acumulado sobre todos os K valores de k e todos os N setores.

Para ilustrar, considere $N = 1$, $K = 4$, com os rankings:

Posição	**Calibração** R_1^4	**Validação** T_1^4
1º	MLP	RNN
2º	RNN	RF
3º	NB	MLP
4º	RF	NB

O vetor $\text{rank}_{RT}[1:k]$ registra, para cada posição k do ranking de validação, qual posição aquele modelo ocupava na calibração:

k	Modelo (valid.)	Posição na calibração	$\text{rank}_{RT}[1:k]$
1	RNN	2º	$[2]$
2	RF	4º	$[2,4]$
3	MLP	1º	$[2,4,1]$
4	NB	3º	$[2,4,1,3]$

Aplica-se então a fórmula para cada k, contando quantos elementos do vetor são menores que k:

k	$\sum I\{\text{rank}_{RT}[1:k] < k\}$	$\sum I + 1$	Termo ($\div k$)
1	0	1	$1/1 = 1{,}0000$
2	0	1	$1/2 = 0{,}5000$
3	2	3	$3/3 = 1{,}0000$
4	3	4	$4/4 = 1{,}0000$

Em $k = 2$, nenhum elemento de $[2{,}4]$ é menor que 2, o que penaliza o fato de que o segundo modelo mais relevante na validação (RF) estava na última posição da calibração. Em $k = 3$, dois elementos de $[2{,}4{,}1]$ são menores que 3 (o 2 e o 1), o que indica que MLP e RNN estavam bem posicionados na calibração. Finalmente:

$$\text{MAP} = \frac{1}{1 \times 4}\left(1{,}0000 + 0{,}5000 + 1{,}0000 + 1{,}0000\right) = \frac{3{,}5}{4} = 0{,}88.$$

O valor de 0,88 é elevado porque a maioria dos modelos relevantes na validação estava bem posicionada na calibração; a única penalização ocorreu em $k = 2$, quando RF (2º na validação) ocupava a última posição na calibração. Um ranking de validação idêntico ao de calibração produziria MAP $= 1{,}0$, enquanto uma ordem completamente invertida aproximaria o MAP de zero;

7. **NDCG@k** (Ganho Cumulativo Descontado Normalizado, do inglês *Normalized Discounted Cumulative Gain*): mede a consistência da ordenação entre os rankings de calibração e validação, penalizando discordâncias em posições mais altas do ranking. Diferentemente das métricas PR, RR e F1, que avaliam apenas *quais* modelos aparecem no top-k sem considerar a ordem, o NDCG penaliza também *trocas de posição* entre modelos, sendo portanto o indicador mais exigente do conjunto.
   O elemento central do cálculo é o **coeficiente de correlação de Spearman** entre dois vetores de tamanho i. Por ser uma medida não paramétrica baseada em postos, é mais adequada para dados ordinais como rankings. Para cada posição i do ranking de validação, identificam-se os top-i modelos da validação e registram-se as posições que esses mesmos modelos ocupavam no ranking de calibração. Isso gera dois vetores:

- **Vetor A**: posições dos modelos no ranking de *validação*, que por construção é sempre $(1, 2, \ldots, i)$;
- **Vetor B**: posições desses mesmos modelos no ranking de *calibração*, que varia conforme a concordância entre os dois rankings.

O coeficiente $c_i = \text{corr}(\mathbf{A}, \mathbf{B}) \in [-1, 1]$ mede o grau de concordância ordinal: $c_i = 1$ indica ordem idêntica; $c_i = 0$ indica ausência de relação linear; $c_i = -1$ indica inversão completa. Para $i = 1$, a correlação é indefinida por ausência de variância e adota-se a convenção: $c_1 = 1$ se o modelo que ocupa a primeira posição na validação também ocupa a primeira posição na calibração, e $c_1 = 0$ caso contrário.

O **ganho cumulativo** (CG) agrega as correlações de todos os setores:

$$\text{CG@}k = \sum_{n=1}^{N} \text{corr}_n(\text{rank}_{RT}[1:k],\, 1:k) = \sum_{n=1}^{N} \text{rel}_n, \tag{83}$$

onde $\text{corr}_n(\text{rank}_{RT}[1:k],\, 1:k)$ é o coeficiente de correlação entre os vetores de tamanho k descritos acima para o n-ésimo setor.

O **ganho cumulativo descontado** (DCG) pondera cada correlação por um fator logarítmico decrescente, de modo que discordâncias nas primeiras posições penalizam mais do que discordâncias nas posições inferiores:

$$\text{DCG}_n\text{@}k = \sum_{i=1}^{k} \frac{2^{\text{corr}_n(\text{rank}_{RT}[1:i],\, 1:i)} - 1}{\log_2(i+1)}. \tag{84}$$

Cálculo detalhado dos coeficientes de correlação. Considere $N = 2$ setores e $k = 4$, com os seguintes rankings:

Setor 1			**Setor 2**		
Posição	Calibração R_1^4	Validação T_1^4	Posição	Calibração R_2^4	Validação T_2^4
1º	MLP	RNN	1º	RNN	RNN
2º	RNN	RF	2º	MLP	RF
3º	NB	MLP	3º	RF	MLP
4º	RF	NB	4º	NB	NB

Para facilitar o cálculo, registra-se a posição de cada modelo em cada ranking:

Setor 1			**Setor 2**		
Modelo	Pos. Calib.	Pos. Valid.	Modelo	Pos. Calib.	Pos. Valid.
RNN	2	1	RNN	1	1
RF	4	2	RF	3	2
MLP	1	3	MLP	2	3
NB	3	4	NB	4	4

Os coeficientes c_i de cada setor são calculados a seguir.

Setor 1 — coeficientes c_1 a c_4:

- $i = 1$: top-1 da validação é RNN, que ocupava a posição 2 na calibração. Como $\text{rank}_{RT}[1] = 2 \neq 1$, aplica-se a convenção $c_1 = 0$.
- $i = 2$: top-2 da validação são RNN e RF. Vetor A $= [1, 2]$; Vetor B $= [2, 4]$. Ambos os vetores são estritamente crescentes, logo a correlação é $c_2 = 1{,}00$.
- $i = 3$: top-3 são RNN, RF, MLP. Vetor A $= [1, 2, 3]$; Vetor B $= [2, 4, 1]$. Calculando:

$$\bar{A} = 2{,}0, \quad \bar{B} = 7/3 \approx 2{,}333;$$

$$\text{num} = (-1)(-0{,}333) + (0)(1{,}667) + (1)(-1{,}333) = 0{,}333 - 1{,}333 = -1{,}000;$$

$$\text{den} = \sqrt{2} \times \sqrt{4{,}667} \approx 3{,}055; \quad c_3 = -1{,}000/3{,}055 \approx -0{,}33.$$

- $i = 4$: top-4 são RNN, RF, MLP, NB. Vetor A $= [1, 2, 3, 4]$; Vetor B $= [2, 4, 1, 3]$. Com $\bar{A} = \bar{B} = 2{,}5$:

$$\text{num} = (-1{,}5)(-0{,}5)+(-0{,}5)(1{,}5)+(0{,}5)(-1{,}5)+(1{,}5)(0{,}5) = 0{,}75-0{,}75-0{,}75+0{,}75 = 0;$$

Aplicando a Equação (84):

i	c_i	$2^{c_i} - 1$	$\log_2(i+1)$	Termo
1	0	0,000	1,000	0,000
2	1	1,000	1,585	0,631
3	−0,33	−0,204	2,000	−0,102
4	0	0,000	2,322	0,000
			$\text{DCG}_1@4$	$\mathbf{0{,}53}$

Setor 2 — coeficientes c_1 a c_4:

- $i = 1$: top-1 é RNN, que ocupava a posição 1 na calibração. Pela convenção, $c_1 = 1$.
- $i = 2$: Vetor A $= [1, 2]$; Vetor B $= [1, 3]$. Ambos crescentes, $c_2 = 1{,}00$.
- $i = 3$: Vetor A $= [1, 2, 3]$; Vetor B $= [1, 3, 2]$. Com $\bar{A} = \bar{B} = 2$:

$$\text{num} = (-1)(-1)+(0)(1)+(1)(0) = 1{,}000; \quad \text{den} = \sqrt{2}\times\sqrt{2} = 2{,}000; \quad c_3 = 0{,}50.$$

- $i = 4$: Vetor A $= [1, 2, 3, 4]$; Vetor B $= [1, 3, 2, 4]$. Com $\bar{A} = \bar{B} = 2{,}5$:

$$\text{num} = (-1{,}5)(-1{,}5)+(-0{,}5)(0{,}5)+(0{,}5)(-0{,}5)+(1{,}5)(1{,}5) = 2{,}25-0{,}25-0{,}25+2{,}25 = 4;$$

$$\text{den} = \sqrt{5} \times \sqrt{5} = 5{,}00; \quad c_4 = 0{,}80.$$

i	c_i	$2^{c_i}-1$	$\log_2(i+1)$	Termo
1	1,00	1,000	1,000	1,000
2	1,00	1,000	1,585	0,631
3	0,50	0,414	2,000	0,207
4	0,80	0,741	2,322	0,319
			$\text{DCG}_2@4$	$\mathbf{2,16}$

O NDCG normaliza o DCG pelo valor ideal (IDCG), que corresponde ao caso em que a ordenação de validação é idêntica à de calibração ($c_i = 1$ em todas as posições):

$$\text{NDCG}@k = \frac{1}{N}\sum_{n=1}^{N}\frac{\text{DCG}_n@k}{\text{IDCG}}, \tag{85}$$

onde $\text{IDCG} = \sum_{i=1}^{k}(2^1-1)/\log_2(i+1)$. Para $k=4$:

i	2^1-1	$\log_2(i+1)$	Termo
1	1	1,000	1,000
2	1	1,585	0,631
3	1	2,000	0,500
4	1	2,322	0,431
		IDCG	$\mathbf{2,56}$

Portanto:

$$\text{NDCG}@4 = \frac{1}{2}\left(\frac{0,53}{2,56}+\frac{2,16}{2,56}\right) = \frac{0,53+2,16}{2\times 2,56} = 0,53.$$

Na adaptação empregada, valores mais altos de NDCG indicam maior consistência entre os rankings de calibração e validação, enquanto valores próximos de zero indicam baixa concordância na ordenação dos modelos. Vale notar que o NDCG é mais exigente do que PR, RR e F1: mesmo que os mesmos modelos apareçam no top-k dos dois rankings (PR e RR altos), o NDCG será penalizado se a ordem entre eles não for preservada.

Um resumo de como funciona cada métrica pode ser conferido na Tabela 11.

5.2 Resultados Experimentais

Esta seção analisa o desempenho dos diferentes modelos de negociação nas granularidades setorial e individual. Em seguida, aplica-se o algoritmo de recomendação proposto para identificar o modelo de negociação mais adequado a cada contexto. Por fim, com base nas recomendações obtidas, constroem-se quatro portfólios *long-short*, ponderados segundo diferentes esquemas de alocação de capital, a fim de avaliar o valor prático do sistema proposto.

Tabela 11 – Forma de agregação dos 8 setores por métrica de generalização.

Métrica	**Como agrega os 8 setores**
PR@k	União dos top-k de todos os setores em um único conjunto; compara-se com a união do cenário de validação.
RR@k	Idem ao PR@k; numerador idêntico, denominador é a cardinalidade da união do cenário de validação.
F1@k	Derivado diretamente de PR@k e RR@k (média harmônica); não exige agregação adicional.
CR@k	Cardinalidade da união dos top-k de todos os setores no cenário de validação, dividida pelo total de modelos candidatos ($K = 12$).
MRR	Calculado setor a setor: para cada setor, obtém-se a posição do modelo campeão da validação no ranking de calibração ($1/\text{rank}$); o valor final é a média das 8 contribuições.
MAP	Calculado setor a setor: para cada setor, computa-se a precisão média ao longo das posições do ranking; o valor final é a média sobre todos os setores e posições.
NDCG@k	Calculado setor a setor: para cada setor, correlaciona-se a ordem dos top-k modelos da validação com as posições que esses modelos ocupavam no ranking de calibração; normaliza-se pelo caso ideal (IDCG) e calcula-se a média dos 8 setores.

Nota: PR, RR, F1 e CR tratam todos os setores como um único *pool* (via união dos conjuntos), enquanto MRR, MAP e NDCG calculam uma medida por setor e depois calculam a média (capturando, portanto, aspectos distintos da qualidade de generalização).

5.2.1 Comparação do Desempenho dos Modelos com e sem Seleção de Atributos por Informação Mútua

Primeiramente, para avaliar o impacto da seleção de atributos sobre o desempenho dos modelos de *trading*, cada um dos 12 modelos candidatos foi treinado em duas configurações: utilizando todos os 44 indicadores técnicos e utilizando o subconjunto de 23 indicadores selecionados pelo fluxo de processamento de Correlação de Pearson e Informação Mútua (Seção 4.1). A comparação foi realizada sobre o cenário InS-TraD (2018–2019), calculando as oito métricas de desempenho para cada ação e obtendo a média geral sobre as 239 ações. A significância estatística das diferenças foi avaliada pelo teste de Wilcoxon pareado (RAMOS; FREIRE; BARBOSA, 2018). As Tabelas 12 a 23 apresentam os resultados de cada modelo. Valores de p-valor inferiores a $0{,}05$ são marcados com *.

Tabela 12 – Comparação NB: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5159	0,5169	0,6815	MI
ARR	0,2893	0,2923	0,3146	MI
ASR	0,8562	0,8826	0,4917	MI
MDD	0,2780	0,2746	0,1677	MI
CAR	1,8079	1,8271	0,1477	MI
SOR	0,0444	0,0456	0,2941	MI
MSR	−0,5281	−0,3970	0,1575	MI
ART	60,88	58,09	0,0339*	MI
MI vence em 8/8 métricas				

Tabela 13 – Comparação LR: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5126	0,5112	0,2859	44
ARR	0,3045	0,3220	0,4409	MI
ASR	0,8314	0,8358	0,8057	MI
MDD	0,3044	0,2974	0,2562	MI
CAR	1,6094	1,6395	0,2641	MI
SOR	0,0454	0,0467	0,3511	MI
MSR	−0,2546	−0,2216	0,7120	MI
ART	65,18	59,69	0,7669	MI
MI vence em 7/8 métricas				

Tabela 14 – Comparação MLP: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,4589	0,4475	0,8702	44
ARR	0,2854	0,3199	0,9421	MI
ASR	0,6444	0,6580	0,9266	MI
MDD	0,3143	0,3075	0,4629	MI
CAR	1,1807	1,2058	0,9369	MI
SOR	0,0434	0,0447	0,9132	MI
MSR	0,1123	0,3275	0,0295*	MI
ART	72,29	68,27	0,0720	MI
MI vence em 7/8 métricas				

A Tabela 24 consolida os resultados dos 12 modelos, ordenados pelo número de métricas em que a seleção por MI apresentou melhor desempenho.

Os resultados revelam um padrão consistente: a seleção de atributos por Informação Mútua beneficia modelos que não possuem mecanismo interno de compressão ou seleção de *features* (NB, LR, MLP, RNN), é marginalmente benéfica para modelos com mecanismo parcial (XGB com apenas 15 rodadas de *boosting*), e é redundante ou prejudicial para modelos que já realizam compressão ou seleção implícita (SAE, SVM, CART, RF, LSTM, GRU). O caso do DBN, que apresentou empate (4/8), é explicado pelo fato de seu pré-treinamento

Tabela 15 – Comparação RNN: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,4929	0,4973	0,1156	MI
ARR	0,2186	0,2351	0,3431	MI
ASR	0,5343	0,5990	0,4061	MI
MDD	0,3332	0,3183	0,1384	MI
CAR	0,9584	1,0922	0,3284	MI
SOR	0,0305	0,0347	0,3557	MI
MSR	0,0983	−0,0012	0,6565	44
ART	101,99	84,80	0,2788	MI
MI vence em 7/8 métricas				

Tabela 16 – Comparação XGB: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5094	0,5100	0,5497	MI
ARR	0,2963	0,3305	0,9189	MI
ASR	0,7958	0,8313	0,7692	MI
MDD	0,3098	0,3111	0,4515	44
CAR	1,4704	1,7631	0,4992	MI
SOR	0,0439	0,0468	0,8789	MI
MSR	−0,2156	−0,3349	0,8598	44
ART	71,56	67,53	0,4285	MI
MI vence em 6/8 métricas				

Tabela 17 – Comparação DBN: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,4529	0,4569	0,5009	MI
ARR	0,3083	0,3335	0,5332	MI
ASR	0,6663	0,6483	0,5170	44
MDD	0,3213	0,3139	0,6831	MI
CAR	1,2177	1,2038	0,5891	44
SOR	0,0437	0,0436	0,4104	44
MSR	0,1051	0,1655	0,3089	MI
ART	67,15	72,43	0,3675	44
Empate: MI vence em 4/8 métricas				

com *autoencoder* oferecer compressão menos eficaz do que a do SAE, motivo pelo qual foi adotada a configuração com seleção de atributos, devido ao melhor desempenho na métrica MSR. O contraste entre RNN (7/8 para MI) e LSTM/GRU (1/8 para MI) é particularmente revelador, pois isola o mecanismo de *gates* como o fator responsável pela seleção implícita de *features* nos modelos recorrentes.

Com base nesses resultados, a configuração adotada para as etapas subsequentes (*backtesting*, TODIM e portfólios) utiliza sinais de MI para NB, DBN, LR, MLP, RNN e XGB, e sinais com 44 *features* para SAE, SVM, CART, RF, LSTM e GRU.

Tabela 18 – Comparação SAE: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,4636	0,4444	0,1797	44
ARR	0,3485	0,3195	0,1554	44
ASR	0,6871	0,6130	0,1740	44
MDD	0,3280	0,3077	0,1477	MI
CAR	1,2853	1,1961	0,1777	44
SOR	0,0463	0,0415	0,1565	44
MSR	0,1327	0,1136	0,0312*	44
ART	71,92	70,76	0,8641	MI
44 features vencem em 6/8 métricas				

Tabela 19 – Comparação SVM: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5131	0,5119	0,3244	44
ARR	0,3785	0,3778	0,8348	44
ASR	0,9106	0,8902	0,6204	44
MDD	0,3055	0,3086	0,3890	44
CAR	1,7922	1,7477	0,3585	44
SOR	0,0532	0,0517	0,5529	44
MSR	−0,2831	−0,2052	0,6503	MI
ART	60,53	68,81	0,0042*	44
44 features vencem em 7/8 métricas				

Tabela 20 – Comparação CART: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5098	0,5086	0,5729	44
ARR	0,3559	0,3313	0,7336	44
ASR	0,8263	0,7964	0,7336	44
MDD	0,3003	0,3124	0,0714	44
CAR	1,6003	1,5511	0,5659	44
SOR	0,0466	0,0451	0,8465	44
MSR	−0,1146	−0,2451	0,1541	44
ART	68,52	62,12	0,8366	MI
44 features vence em 7/8 métricas				

5.2.2 Análise de Desempenho de Algoritmo de Negociação de Ações Setoriais

Para melhorar a leitura deste trabalho, seleciona-se apenas o *Sharpe Ratio Anualizado* (ASR) dos dados In-Sample de treino e de teste, como resultado da pesquisa, para exibição e discussão. Além disso, calcula-se o desempenho da estratégia de comprar e segurar (do inglês *buy and hold*, BAH) e do índice de mercado[4]. Considera-se que, se o

[4] Utilizou-se o Ibovespa (*Índice Bovespa*), principal indicador de desempenho das ações negociadas na B3, composto pelas empresas de maior liquidez e representatividade do mercado acionário brasileiro.

Tabela 21 – Comparação RF: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5164	0,5161	0,9313	44
ARR	0,3569	0,3496	0,3476	44
ASR	0,9284	0,9184	0,8354	44
MDD	0,2959	0,2998	0,3937	44
CAR	1,8193	1,7514	0,4983	44
SOR	0,0513	0,0509	0,7387	44
MSR	−0,5463	−0,2601	0,0829	MI
ART	56,64	57,23	0,9163	44
44 features vence em 7/8 métricas				

Tabela 22 – Comparação LSTM: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5078	0,5055	0,0490*	44
ARR	0,3645	0,3451	0,6486	44
ASR	0,7967	0,7650	0,5139	44
MDD	0,3284	0,3314	0,8723	44
CAR	1,4525	1,3878	0,3780	44
SOR	0,0475	0,0450	0,5676	44
MSR	−0,0964	0,0373	0,5022	MI
ART	71,00	72,79	0,6677	44
44 features vence em 7/8 métricas				

Tabela 23 – Comparação GRU: 44 features vs MI features.

Métrica	44 feat	MI feat	p-valor	Melhor
WR	0,5050	0,5031	0,0175*	44
ARR	0,3420	0,3354	0,3480	44
ASR	0,7497	0,7055	0,1692	44
MDD	0,3315	0,3379	0,1151	44
CAR	1,4025	1,2885	0,1580	44
SOR	0,0449	0,0420	0,2337	44
MSR	−0,0379	0,0221	0,6862	MI
ART	77,13	83,18	0,0613	44
44 features vence em 7/8 métricas				

desempenho do algoritmo de negociação for pelo menos superior ao do BAH ou ao do índice de mercado, então esse algoritmo possui valor de aplicação.

Na Tabela 25, o melhor desempenho de ASR em cada setor é destacado em negrito. Nas instâncias em que as estratégias de referência (*benchmarks*) superam todos os modelos de aprendizado de máquina, seus valores também são destacados em negrito; nesses casos, o modelo com a melhor performance relativa é sinalizado com um asterisco (*). A análise dos resultados demonstra que o Índice de Mercado (Ibovespa) supera todos os algoritmos de ML na maioria dos setores, o que é consistente com a hipótese de maior

Tabela 24 – Resumo consolidado: número de métricas vencidas pela configuração MI em cada modelo.

Modelo	**Mecanismo interno**	**MI vence**	**Configuração adotada**
NB	Nenhum (assume independência)	8/8	MI
LR	Nenhum	7/8	MI
MLP	Nenhum	7/8	MI
RNN	Nenhum (sem *gates*)	7/8	MI
XGB	*Boosting* (15 rodadas)	6/8	MI
DBN	Pré-treino *autoencoder*	4/8	MI
SAE	*Autoencoder*	2/8	44 features
SVM	Kernel RBF	1/8	44 features
CART	Seleção por *split*	1/8	44 features
RF	Ensemble (500 árvores)	1/8	44 features
LSTM	*Gates* (forget/input/output)	1/8	44 features
GRU	*Gates* (reset/update)	1/8	44 features

eficiência do mercado brasileiro no período amostral. A exceção ocorre no setor FIN, em que o RF obtém ASR de 1,3148, superando o índice (1,2991) e o BAH (1,1156), o que demonstra que há valor na negociação ativa nesse setor. No setor UTI, o BAH supera todos os algoritmos, sendo NB o mais próximo da estratégia passiva, com ASR de 1,4060 frente a 1,4773 do BAH. Em todos os demais setores, embora os modelos não superem o índice de mercado, o melhor algoritmo de cada setor apresenta desempenho superior ao BAH, evidenciando que a negociação ativa agrega valor em relação à estratégia passiva de manutenção de posição[5].

A Tabela 26 apresenta os valores de ASR para o período de teste das ações In-Sample (InS-TesD), correspondente ao primeiro semestre de 2020. Este período é marcado pelo colapso abrupto dos mercados financeiros globais em decorrência da pandemia de COVID-19, o que confere a este conjunto de resultados um caráter de teste sob condições extremas de estresse de mercado.

O aspecto mais notável dos resultados é que o Ibovespa registra ASR negativo em todos os oito setores analisados, reflexo direto da queda severa e generalizada observada no mercado brasileiro nesse intervalo. Nesse contexto adverso, os modelos de ML demonstram capacidade de proteção: em todos os setores, o melhor algoritmo supera tanto o Índice de Mercado quanto a estratégia BAH, o que é especialmente relevante, pois evidencia a utilidade da negociação ativa em cenários de alta volatilidade e tendência de baixa.

Em termos absolutos, o setor COM se destaca pelo melhor desempenho individual, com o SVM atingindo um ASR de 1,6457. Os setores GCS e CNC também apresentam resultados expressivos, com NB (0,8047) e MLP (0,7357) como os melhores algoritmos,

[5] Considerando 8 métricas e 4 conjuntos de dados (InS-TraD, InS-TesD, OutS-TraD e OutS-TesD), totalizam-se 32 tabelas, o que torna inviável incluí-las todas no corpo do texto. As tabelas faltantes podem ser conferidas no Apêndice C.

Tabela 25 – Valores de *Sharpe Ratio* Anualizado (ASR) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**1,0847**	**1,1253**	**1,1170**	**1,3402**	1,2991	**1,2811**	**1,0951**	1,2406
BAH	0,6605	0,7991	0,4950	0,3861	1,1156	0,7582	0,6866	**1,4773**
LR	0,5267	0,7613	0,6105	0,0082	1,2436	0,6344	0,7261	1,3503
SVM	0,5592	**0,9603***	**0,7064***	0,4304	1,2597	0,6703	**0,9773***	1,3415
NB	0,7137	0,8349	0,6492	0,4725	1,2061	0,5754	0,9053	1,4060*
CART	0,5574	0,8172	0,6140	0,3782	1,0745	**0,8207***	0,6661	1,1676
RF	**0,7629***	0,8231	0,6496	0,4671	**1,3148**	0,6794	0,9505	1,3649
XGB	0,5908	0,8232	0,5389	0,2662	1,0780	0,7755	0,2508	1,3083
MLP	0,4080	0,6324	0,2969	0,3873	0,8474	0,6707	0,3401	1,0472
DBN	0,3391	0,6831	0,5777	0,1468	0,9578	0,5968	0,3576	1,1630
SAE	0,4358	0,5603	0,5195	0,4722	0,9947	0,6314	0,1295	1,1992
RNN	0,5559	0,6533	0,4461	0,5904	0,9968	0,5813	0,7308	1,3198
LSTM	0,6446	0,7416	0,4165	**0,6227***	0,9729	0,5915	0,3722	1,2086
GRU	0,5850	0,6475	0,2371	0,2756	0,9768	0,5187	0,4091	1,3262

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) indica o melhor modelo de ML quando os *benchmarks* superam todos os algoritmos. Em COM, tanto RNN quanto LSTM apresentam os dois maiores valores entre os algoritmos de ML, sendo LSTM o de maior ASR.

respectivamente. No setor FIN, embora o melhor modelo (GRU, $-0{,}0559$) ainda apresente ASR ligeiramente negativo, ele supera com folga tanto o Índice ($-0{,}5071$) quanto o BAH ($-0{,}3667$), demonstrando que a estratégia ativa amorteceu significativamente as perdas nesse setor. Vale destacar o protagonismo das redes recorrentes nesse período: GRU é o melhor modelo em quatro setores (BM, CC, FIN e OGB), e SVM lidera em COM e UTI, sugerindo que modelos com maior capacidade de capturar dependências temporais ou margens de separação robustas se adaptam melhor a regimes de mercado em crise.

5.2.3 O Modelo de Negociação ótimo para Ações Setoriais e Individuais

Com base nos resultados da análise acima (incluindo as demais métricas), obtém-se o modelo de negociação ideal para investimento em ações, com base no algoritmo de recomendação TODIM, aplicado em duas granularidades.

Na abordagem setorial, a matriz de decisão de cada setor é construída a partir da média aritmética das métricas de todas as ações desse setor. Ou seja, para cada modelo candidato e cada métrica, calcula-se a média dos valores obtidos em todas as ações do setor, resultando em uma matriz de decisão de 12×8 (12 modelos candidatos $\times$ 8 métricas). O TODIM é então aplicado a essa matriz para gerar um ranking de modelos por setor, identificando qual apresenta o melhor desempenho médio em cada um.

Na abordagem individual, as métricas de cada ação são utilizadas diretamente, sem agregação. Para cada ação, constrói-se uma matriz de decisão 12×8 com os valores das

Tabela 26 – Valores de *Sharpe Ratio* Anualizado (ASR) no cenário InS-TesD

Modelo	BM	CC	CNC	COM	FIN	GCS	OGB	UTI
Index	−0,6085	−0,6955	−0,4753	−0,4748	−0,5071	−0,4282	−0,5119	−0,4543
BAH	−0,0179	−0,0162	0,4515	0,8055	−0,3667	0,1653	0,1048	0,0333
LR	0,2292	0,1397	0,4744	0,3923	−0,4595	0,4578	−0,4403	0,2145
SVM	0,3057	0,0523	0,2477	**1,6457**	−0,4496	0,5608	0,2915	**0,5703**
NB	0,1663	0,0782	−0,3325	1,1697	−0,2303	**0,8047**	0,1469	0,0313
CART	0,0139	0,1154	0,7143	−0,0727	−0,2077	−0,0272	−0,1501	−0,0062
RF	−0,1247	−0,0788	−0,1981	1,1842	−0,3816	0,2915	0,2839	0,2862
XGB	0,1209	0,0983	−0,0696	0,2078	−0,2105	0,0568	0,3391	0,0497
MLP	−0,1536	0,0275	**0,7357**	1,0613	−0,4062	0,1456	−0,3267	−0,0843
DBN	−0,2831	−0,0006	0,7130	0,6207	−0,3782	0,2916	0,0763	0,0007
SAE	−0,1956	−0,1084	0,6541	0,7557	−0,3779	0,1975	−0,0175	−0,0173
RNN	0,2488	0,1568	0,3923	0,9317	−0,1506	0,2075	0,2628	0,2832
LSTM	0,1243	−0,1118	0,5376	0,3193	−0,4950	0,0274	0,0266	0,1984
GRU	**0,3433**	**0,1604**	0,0049	0,5270	**−0,0559**	0,4821	**0,5675**	0,2972

Nota: Valores em **negrito** indicam o melhor desempenho na coluna (algoritmo de ML que superou ambos os *benchmarks* em todos os setores). Valores negativos do Índice e do BAH refletem o impacto da pandemia de COVID-19 no primeiro semestre de 2020.

métricas da ação específica, e o TODIM é aplicado separadamente para identificar o modelo ótimo para a ação. Essa abordagem visa mostrar que ações de um mesmo setor podem apresentar dinâmicas distintas e, portanto, se beneficiar de modelos distintos.

A importância relativa dos critérios de avaliação é crucial no algoritmo de recomendação em ambas as granularidades. A seguir, apresentam-se as razões que justificam a ordem de importância adotada. WR é um indicador utilizado para avaliar a precisão das recomendações de negociação (i.e., os *trading signals*, TSs), ou seja, a proporção de TSs corretos em relação ao total de TSs. No entanto, um WR elevado não conduz necessariamente a um lucro elevado. MDD e ART são indicadores de risco. O MDD calcula a profundidade máxima do *drawdown*, enquanto o ART estima a duração média do período de recuperação dos *drawdowns* observados. ARR é um indicador de rentabilidade. O objetivo da negociação de ações é obter lucro; portanto, o ARR é um índice de avaliação mais crítico do que o MDD e o ART. ASR, SOR e CAR consideram tanto o retorno quanto o risco do modelo de negociação, porém, as funções de risco desses três indicadores são distintas: variância, semivariância e MDD. A variância mede o quanto os retornos se desviam da média. No entanto, o desvio dos retornos positivos em relação à média também é considerado um risco. Dessa forma, o SOR constitui uma métrica de avaliação mais adequada do que o ASR. A função de risco do CAR é o MDD. Se a semivariância considera o risco de a série de retornos ser inferior à sua média, o MDD considera o risco da perda máxima possível. Na negociação de ações, é necessário observar atentamente a perda máxima potencial. Portanto, neste trabalho, assim como em Lv et al. (2023), considera-se o CAR uma métrica mais crítica do que o SOR. Além disso, o MSR considera retornos, riscos e a distribuição

dos retornos, isto é, assimetria e curtose, visto que séries temporais financeiras apresentam tipicamente distribuições com picos acentuados e caudas pesadas, em vez de distribuições normais. Portanto, é válido afirmar que o MSR é o mais importante entre os oito indicadores de avaliação de desempenho de negociação. A partir da análise acima, conclui-se que:

$$q_{\text{MSR}} > q_{\text{CAR}} > q_{\text{SOR}} > q_{\text{ASR}} > q_{\text{ARR}} > q_{\text{MDD}} = q_{\text{ART}} > q_{\text{WR}}.$$

Se $q_1 = q_{\text{WR}} = 1$, então define-se $q_2 = q_{\text{MDD}} = q_3 = q_{\text{ART}} = 2$, $q_4 = q_{\text{ARR}} = 3$, $q_5 = q_{\text{ASR}} = 4$, $q_6 = q_{\text{SOR}} = 5$, $q_7 = q_{\text{CAR}} = 6$ e $q_8 = q_{\text{MSR}} = 7$. Portanto, a ordem de importância dos indicadores de avaliação é MSR $>$ CAR $>$ SOR $>$ ASR $>$ ARR $>$ MDD $=$ ART $>$ WR, e sua importância é definida por $q = (7, 6, 5, 4, 3, 2, 2, 1)$. Esses pesos são idênticos em ambas as granularidades (setorial e individual), garantindo que a comparação entre as duas abordagens reflita exclusivamente o efeito da agregação, e não diferenças nos critérios de avaliação. Os resultados experimentais fornecem o modelo ótimo (top-1) para cada setor e para cada ação individual.

5.2.4 Modelo Ótimo de Negociação de Ações Setoriais

Com base nos resultados de desempenho apresentados anteriormente, aplica-se o algoritmo de recomendação TODIM para obter o modelo ótimo de negociação de ações setoriais para cada um dos quatro cenários avaliados (InS-TraD, InS-TesD, OutS-TraD e OutS-TesD). A Tabela 27 apresenta o modelo top-1[6] recomendado por setor em cada cenário.

Tabela 27 – Modelo de negociação top-1 recomendado pelo TODIM por setor e cenário.

Setor	InS-TraD	InS-TesD	OutS-TraD	OutS-TesD
BM	RF	SVM	LR	NB
CC	SVM	LR	LR	XGB
CNC	NB	CART	RNN	GRU
COM	RNN	NB	CART	XGB
FIN	RF	CART	SVM	RNN
GCS	CART	SVM	RF	DBN
OGB	NB	GRU	CART	XGB
UTI	GRU	SVM	RF	GRU

Uma primeira observação relevante é a ausência de um modelo universalmente dominante: ao longo dos 32 slots (8 setores $\times$ 4 cenários), nenhum algoritmo se impõe de forma consistente. Os modelos SVM e CART lideram com cinco aparições cada no top-1, seguidos por RF, NB e GRU (quatro cada), RNN e LR (três cada) e XGB (três). Esse resultado está alinhado com a literatura, que aponta a dependência do algoritmo ótimo das características específicas de cada setor e de cada período (LV et al., 2023).

[6] Vale lembrar que, o TODIM constrói um ranking para cada setor/ação individual, logo, existe um top-12, todavia, por uma questão de economia de espaço se optou por mostrar apenas o top-1.

Em termos de consistência intrasetorial, destacam-se dois casos: no setor **CC**, o LR ocupa o top-1 nos cenários InS-TesD e OutS-TraD, sugerindo estabilidade relativa desse algoritmo no segmento de consumo cíclico; no setor **UTI**, o GRU lidera tanto em InS-TraD quanto em OutS-TesD, indicando que redes recorrentes do tipo *Gated Recurrent Unit* capturam padrões mais persistentes nas ações do setor de utilidades. Nos demais setores, a alternância de modelos entre cenários é a regra, o que reforça a importância do algoritmo de recomendação multicritério: não é possível eleger um único modelo *a priori* sem considerar o conjunto de indicadores de desempenho e o cenário de aplicação.

Merece destaque também a frequência com que modelos clássicos de ML figuram no top-1. Algoritmos como RF, SVM, NB, CART, LR e XGB aparecem em 24 dos 32 slots, enquanto modelos neurais ou de aprendizado profundo (RNN, GRU e DBN) ocupam os 8 restantes. Esse padrão é consistente com a análise de ações individuais discutida na Seção 5.2.5 e com a literatura, que demonstra que algoritmos tradicionais podem ser tão ou mais eficazes do que redes profundas em horizontes curtos de negociação (LV et al., 2023).

5.2.5 Melhor Modelo por Ação Individual

Além da recomendação setorial, obteve-se o melhor modelo de negociação para cada ação individual, considerando os períodos de treino (InS-TraD e OutS-TraD). Ao todo, foram analisadas 239 ações, sendo 190 *In-Sample* e 49 *Out-of-Sample*, distribuídas entre os oito setores da B3ICS. Vale ressaltar que as métricas de generalização foram calculadas exclusivamente para a análise de recomendação setorial, não sendo estendidas ao nível de ação individual, uma vez que o objetivo dessas métricas é avaliar a capacidade de transferência do modelo ótimo do conjunto *In-Sample* para o *Out-of-Sample* no contexto agregado de cada setor.

A Tabela 28 apresenta o modelo vencedor por ação no período de treino. Entre as ações *In-Sample*, os cinco algoritmos mais frequentemente recomendados são RF (27 ações; 14,2%), SVM (26; 13,7%), XGB (24; 12,6%), NB (23; 12,1%) e LR (20; 10,5%), totalizando 63,2% das recomendações. Considerando também o CART, os modelos clássicos respondem por uma parcela ainda maior das indicações, reforçando sua predominância no conjunto *In-Sample*.

Os modelos neurais ou de aprendizado profundo, embora presentes, têm participação menor: LSTM (13), RNN (11), MLP (10), SAE (7), GRU (6) e DBN (5), representando coletivamente 27,4% das ações.

O padrão é semelhante nas ações *Out-of-Sample*, em que SVM (7), RF (6), NB (6), LR (5), CART (5) e XGB (4) predominam, com os modelos tradicionais respondendo por 67,3% das indicações.

Tabela 28 – Melhor modelo de negociação por ação individual (período de treino).

Ação	**Modelo**	**Ação**	**Modelo**	**Ação**	**Modelo**	**Ação**	**Modelo**
BRAP3	XGB	BRAP4	LR	BRKM3	LR	BRKM5	XGB
CRPG5	SVM	CRPG6	CART	CSNA3	LSTM	DEXP3	RF
DXCO3	XGB	EUCA4	SVM	FESA3	XGB	FESA4	LSTM
FHER3	MLP	GGBR3	NB	GGBR4	NB	GOAU3	LSTM
GOAU4	RF	KLBN11	SAE	KLBN3	LSTM	KLBN4	RF
MGEL4	GRU	PMAM3	RF	RANI3	GRU	SNSY5	NB
SUZB3	SVM	AALR3	NB	ALPA3	XGB	ALPA4	NB
AMAR3	DBN	AMER3	RNN	ANIM3	XGB	AZZA3	XGB
BHIA3	RF	BIOM3	XGB	BMKS3	CART	CEDO4	RF
CGRA3	RF	CGRA4	NB	COGN3	NB	CTKA4	SVM
CTSA4	SVM	CVCB3	XGB	CYRE3	RNN	DASA3	SVM
DIRR3	RNN	DOHL4	MLP	ESTR4	RF	EVEN3	CART
EZTC3	LSTM	FLRY3	NB	GFSA3	RNN	GRND3	DBN
HBOR3	CART	HETA4	NB	HOOT4	LR	HYPE3	NB
JFEN3	CART	JHSF3	DBN	LEVE3	XGB	LREN3	SVM
MEAL3	NB	MGLU3	SAE	MNDL3	XGB	MOVI3	MLP
MRVE3	RF	MYPK3	SVM	ODPV3	LR	OFSA3	NB
PFRM3	NB	PNVL3	SVM	PTNT4	SVM	QUAL3	LR
RADL3	SVM	RENT3	SVM	RSID3	SVM	ABEV3	LSTM
AGRO3	XGB	BAUH4	SVM	BEEF3	SVM	BOBR4	LR
CAML3	NB	FICT3	CART	MBRF3	XGB	MDIA3	DBN
MNPR3	CART	OIBR3	GRU	OIBR4	LSTM	TELB3	MLP
TELB4	RNN	TIMS3	NB	ABCB4	SVM	B3SA3	RF
BAZA3	RNN	BBAS3	SVM	BBDC3	SVM	BBDC4	SAE
BBSE3	XGB	BEES3	XGB	BEES4	RF	BGIP4	RF
BMEB3	RF	BMEB4	RF	BMIN4	RF	BNBR3	MLP
BPAC11	NB	BPAC3	LR	BPAC5	XGB	BRSR3	LR
BRSR6	SVM	CSUD3	NB	GSHP3	CART	IGTI3	LR
IRBR3	LSTM	ITSA3	XGB	ITSA4	SVM	ITUB3	LR
ITUB4	CART	LPSB3	NB	MERC4	RF	MULT3	XGB
NEXP3	MLP	PINE4	RF	AZEV4	RNN	BDLL4	XGB
DTCY3	RF	EALT3	CART	EALT4	XGB	ECOR3	CART
EPAR3	XGB	ETER3	CART	FRAS3	LR	HAGA4	CART
INEP3	RF	INEP4	SVM	KEPL3	SVM	LOGN3	SAE

Continua na próxima página

Tabela 28 – continuação

Ação	**Modelo**	**Ação**	**Modelo**	**Ação**	**Modelo**	**Ação**	**Modelo**
MILS3	RF	MTSA4	LR	PDTC3	CART	POMO3	LR
POMO4	LSTM	POSI3	LSTM	PSVM11	SVM	PTBL3	LSTM
RAIL3	SVM	RAPT3	CART	RAPT4	LR	RCSL3	GRU
RCSL4	LR	ROMI3	RNN	SHUL4	NB	TASA3	RF
CSAN3	LR	LUPA3	RNN	OSXB3	RF	PETR3	RF
PETR4	NB	PRIO3	RF	RPMG3	SVM	ALUP11	NB
ALUP3	SVM	ALUP4	CART	AXIA3	LSTM	AXIA6	LSTM
CEBR6	RF	CEEB3	RF	CGAS3	LR	CGAS5	SAE
CLSC4	LR	CMIG3	SAE	CMIG4	CART	COCE5	NB
CPFE3	MLP	CPLE3	MLP	CSMG3	GRU	EGIE3	SAE
EKTR4	CART	EMAE4	GRU	ENEV3	XGB	ENGI11	RNN
ENGI3	RF	ENGI4	XGB	ENMT3	LR	EQPA3	LR
EQTL3	DBN	GEPA4	NB	ISAE3	XGB	ISAE4	RNN
LIGT3	MLP	REDE3	MLP				
Ações Out-of-Sample (†):							
UNIP3†	XGB	UNIP5†	LR	UNIP6†	GRU	USIM3†	LR
USIM5†	LR	VALE3†	RF	SEER3†	NB	SHOW3†	DBN
TCSA3†	RF	TECN3†	LSTM	TEND3†	CART	TRIS3†	MLP
UCAS3†	XGB	VIVR3†	LR	VSTE3†	SVM	VULC3†	CART
WHRL3†	MLP	WHRL4†	RF	YDUQ3†	LR	NATU3†	NB
SLCE3†	LSTM	SMTO3†	RNN	VIVT3†	CART	PPLA11†	LSTM
PSSA3†	MLP	SANB11†	RF	SANB3†	RF	SANB4†	SAE
SCAR3†	SVM	SYNE3†	MLP	WIZC3†	NB	TASA4†	SAE
TGMA3†	XGB	TOTS3†	GRU	TPIS3†	RF	TUPY3†	SVM
VLID3†	NB	WEGE3†	SVM	WLMM4†	CART	UGPA3†	LSTM
VBBR3†	XGB	RNEW3†	RNN	RNEW4†	SVM	SAPR3†	CART
SAPR4†	NB	SBSP3†	MLP	TAEE11†	NB	TAEE3†	SVM
TAEE4†	SVM						

Nota: Ações marcadas com † pertencem ao grupo Out-of-Sample.

5.2.6 Avaliação da Capacidade de Generalização do Algoritmo de Recomendação

Primeiramente, constrói-se um modelo de recomendação baseado no InS-TraD e no algoritmo de recomendação proposto na Seção 4.4. Em seguida, constroem-se modelos correspondentes para InS-TesD, OutS-TraD e OutS-TesD. Se os resultados desses três

últimos conjuntos estiverem próximos aos do primeiro, conclui-se que o modelo possui boa capacidade de generalização.

Utilizam-se as métricas da Seção 5.1.5 para calcular os valores de efeito recomendados para esses três conjuntos de dados, conforme apresentado na Tabela 29.

Tabela 29 – Resultados das métricas de generalização por cenário e valor de k.

		Métricas Top-k				**Métricas Globais**		
Cenário	k	**PR**	**RR**	**F1**	**CR**	**MRR**	**MAP**	**NDCG**
InS-TesD	1	0,6667	0,8000	0,7273	0,4167	0,2939	0,6091	0,0000
	3	0,8889	0,8000	0,8421	0,8333			0,3302
	5	0,9091	0,9091	0,9091	0,9167			0,2954
OutS-TraD	1	0,6667	0,8000	0,7273	0,4167	0,2418	0,5942	0,0000
	3	0,8889	1,0000	0,9412	0,6667			0,0818
	5	0,8182	0,9000	0,8571	0,8333			0,1594
OutS-TesD	1	0,5000	0,6000	0,5455	0,4167	0,2441	0,5332	0,1250
	3	0,7778	0,7000	0,7368	0,8333			0,0489
	5	0,9091	0,9091	0,9091	0,9167			0,0304

Nota: PR: *Precision Rate*; RR: *Recall Rate*; F1: média harmônica de PR e RR; CR: *Coverage Rate*; NDCG: *Normalized Discounted Cumulative Gain*; MRR (*Mean Reciprocal Rank*) e MAP (*Mean Average Precision*) são calculados globalmente para o cenário.

Conforme apresentado na Tabela 29, as métricas PR@k, RR@k, F1@k e CR@k crescem de forma aproximadamente monotônica com k, o que é esperado: quanto maior o valor de k, maior a chance de sobreposição entre os top-k modelos recomendados nos conjuntos de validação e os top-k do conjunto de referência (InS-TraD). Em $k = 5$, os três cenários convergem para valores idênticos de PR, RR, F1 e CR (0,9091 para as três primeiras e 0,9167 para CR), indicando alta consistência dos modelos recomendados quando se considera um conjunto mais amplo de candidatos.

O cenário **InS-TesD** apresenta o melhor desempenho global, com MAP de 0,6091 e MRR de 0,2939, indicando que o ranking aprendido no período de treino In-Sample mantém boa utilidade preditiva no período de teste do mesmo conjunto de ações. O cenário **OutS-TraD** exibe RR@3 = 1,0000, ou seja, todos os modelos recomendados para o conjunto Out-of-Sample no período de treino estão no top-3 do ranking In-Sample, o que evidencia forte sobreposição entre as preferências setoriais dos dois grupos de ações. Por outro lado, o cenário **OutS-TesD** apresenta os valores mais baixos em $k = 1$ (PR = 0,5000; F1 = 0,5455), o que é esperado dado que combina a distância amostral (*out-of-sample*) com a distância temporal (período de teste), tornando a tarefa de generalização mais exigente.

Em relação ao NDCG, observa-se um padrão distinto: InS-TesD atinge 0,3302 em $k = 3$ e 0,2954 em $k = 5$, enquanto OutS-TesD apresenta o maior NDCG@1 (0,1250), mas

apresenta valores decrescentes para k maiores. Esse comportamento indica que a *ordem* das recomendações varia mais entre os conjuntos do que no conjunto de modelos em si, aspecto já apontado na literatura como um desafio inerente a algoritmos de recomendação de modelos de *trading* (LV et al., 2023).

Comparando com os resultados reportados por Lv et al. (2023) para os mercados americano e chinês, os valores de MAP (0,5332–0,6091) estão consistentes com a faixa de 0,53–0,76 obtida pelos autores, e o MRR (0,2418–0,2939) está dentro da faixa de 0,18–0,32 relatada. Esses resultados validam a aplicabilidade da metodologia TODIM ao mercado brasileiro, ainda que as características estruturais da B3 (maior concentração setorial, menor liquidez em segmentos específicos e dinâmica macroeconômica distinta) introduzam variabilidade adicional na capacidade de generalização entre períodos e amostras.

5.2.7 Análise de Desempenho de Portfólio Baseado no Algoritmo de Recomendação Dentro da Amostra

A fim de verificar a capacidade de generalização do algoritmo de recomendação na amostra, esta seção apresenta a construção e a avaliação de portfólios baseados no ISRA (*In-Sample Recommendation Algorithm*). Antes de detalhar os resultados, apresentam-se os conceitos fundamentais para a correta interpretação das estratégias adotadas.

Conceitos fundamentais

O que é o ISRA? O ISRA é a estrutura (*framework*) central proposta para selecionar automaticamente o melhor algoritmo de *trading* para cada segmento do mercado. Neste trabalho, amplia-se o escopo original do artigo de referência (LV et al., 2023) ao implementar **duas versões** do ISRA:

- **ISRA_Setor**: replica a abordagem original, selecionando o modelo ótimo por setor (p.ex., RF para o setor FIN, GRU para o setor UTI) por meio do TODIM aplicado às médias setoriais no InS-TraD. O retorno diário do ISRA_Setor é a média aritmética simples dos retornos gerados pelos modelos recomendados em cada um dos 8 setores, pressupondo alocação equitativa de capital entre os segmentos.
- **ISRA_Ind**: contribuição deste trabalho, que estende a recomendação ao nível de ação individual. Para cada uma das 239 ações analisadas, aplica-se o TODIM às métricas de desempenho no período de treino correspondente (InS-TraD para as 190 ações In-Sample; OutS-TraD para as 49 ações Out-of-Sample)[7], selecionando o

[7] Vale destacar uma distinção fundamental entre as duas granularidades de recomendação. No ISRA_Setor, os oito setores são os mesmos em todos os quatro cenários, de modo que o modelo recomendado no InS-TraD pode ser aplicado diretamente tanto ao InS-TesD quanto ao OutS-TesD. No ISRA_Ind, isso não é possível: as 190 ações In-Sample e as 49 ações Out-of-Sample são conjuntos disjuntos, com históricos de sinais e métricas próprios. Assim, o modelo ótimo por ação deve necessariamente ser calibrado no período de treino do grupo ao qual a ação pertence (InS-TraD para as ações In-Sample e OutS-TraD para

modelo ótimo por ativo. O retorno diário do ISRA_Ind é então a média dos retornos de negociação de cada ação operada pelo seu respectivo modelo ótimo individual.

O que é a estratégia Long-Short? A estratégia *long-short* consiste em uma operação que combina simultaneamente uma posição comprada (*long*) em um ativo com expectativa de valorização e uma posição vendida (*short*) em outro ativo com perspectiva de desempenho inferior (HOU et al., 2020). O retorno desta operação combinada é denominado *spread*: a diferença diária entre os retornos dos dois ativos. Por exemplo, o *spread* "ISRA_Ind $-$ BAH" representa, a cada dia, o quanto a estratégia de recomendação individual supera (ou fica aquém) da estratégia passiva de comprar e manter. O objetivo central não é apostar na direção do mercado, mas sim lucrar com a *diferença de desempenho* entre duas estratégias, buscando neutralizar a exposição ao risco sistêmico.

O que é o Índice de Sharpe e como se selecionam os spreads? O Índice de Sharpe (IS) mede o retorno obtido por unidade de risco assumido:

$$\text{IS} = \frac{\bar{r} - r_f}{\sigma_r} \cdot \sqrt{T}, \tag{86}$$

onde $\bar{r}$ é o retorno médio diário do *spread*, r_f é o retorno do ativo livre de risco (assumido como igual a 0[8]), σ_r é seu desvio-padrão e $T = 250$ é o fator de anualização. Um IS > 0 indica que o *spread* gera retorno positivo ajustado ao risco, sendo, portanto, um candidato viável para compor o portfólio. A partir das quatro estratégias disponíveis no período de calibração InS-TraD — Índice, BAH, ISRA_Setor e ISRA_Ind —, formam-se 12 *spreads* possíveis (todas as combinações ordenadas de pares de estratégias). Apenas os *spreads* com IS > 0 no período de calibração são selecionados para compor os portfólios, evitando-se o viés de inclusão de ativos com desempenho historicamente negativo.

Como são calculados os pesos? Após a seleção dos *spreads*, constroem-se quatro portfólios com critérios distintos de alocação de capital:

- **EWP** (*Equal Weight Portfolio*): pesos iguais entre todos os *spreads* selecionados ($w_i = 1/n$).
- **ASRWP** (*ASR Weighted Portfolio*): pesos proporcionais ao Índice de Sharpe de cada *spread* no período de calibração ($w_i \propto \text{IS}_i$), favorecendo os *spreads* com melhor razão retorno-risco histórica.
- **KWP** (*Kelly Weighted Portfolio*) (THORP, 2011): pesos determinados pelo Critério

as ações Out-of-Sample) sendo avaliado no período de teste correspondente (InS-TesD e OutS-TesD, respectivamente). Não seria metodologicamente válido aplicar ao OutS-TesD os modelos calibrados no InS-TraD, pois esses modelos nunca "viram" as ações Out-of-Sample durante a etapa de recomendação.

[8] A adoção de uma taxa livre de risco igual a zero justifica-se pela natureza autofinanciada das estratégias *long-short*, em que a venda a descoberto financia a ponta comprada, e pelo foco do estudo no ranqueamento relativo do poder preditivo dos modelos, visto que a inclusão de uma taxa livre de risco alteraria os valores absolutos do Índice de Sharpe, mas preservaria a ordinalidade dos resultados para fins de seleção de modelos.

de Kelly, que maximiza o crescimento logarítmico esperado do capital. Na prática, resolve-se o sistema que iguala a derivada do retorno esperado logarítmico a zero, o que resulta em alocações mais concentradas nos *spreads* de maior IS.

- **RPWP** (*Risk Parity Weighted Portfolio*) (QIAN, 2005): pesos inversamente proporcionais à volatilidade de cada *spread* ($w_i \propto 1/\sigma_i$), de modo que cada componente contribui igualmente para o risco total do portfólio.

A partir das quatro estratégias disponíveis, calculam-se os Índices de Sharpe (IS) dos 12 *spreads* possíveis, conforme apresentado na Tabela 30. O critério de seleção adotado é IS > 0 no cenário de calibração; seis *spreads* satisfazem essa condição e são utilizados na composição dos portfólios.

Tabela 30 – Índice de Sharpe (IS) dos 12 *spreads* no cenário de calibração (InS-TraD).

Spread	**IS**	**Selecionado**
ISRA_Ind − ISRA_Setor	3,6567	✓
ISRA_Ind − Índice	1,1814	✓
BAH − ISRA_Setor	1,0520	✓
BAH − Índice	1,0204	✓
ISRA_Ind − BAH	0,6766	✓
ISRA_Setor − Índice	0,3511	✓
Índice − ISRA_Setor	−0,3511	
BAH − ISRA_Ind	−0,6766	
Índice − BAH	−1,0204	
ISRA_Setor − BAH	−1,0520	
Índice − ISRA_Ind	−1,1814	
ISRA_Setor − ISRA_Ind	−3,6567	

Nota: Os *spreads* abaixo da linha divisória apresentam IS ≤ 0 e foram excluídos da composição dos portfólios. Por construção, cada par de estratégias gera dois *spreads* simétricos (e.g., ISRA_Ind − ISRA_Setor e ISRA_Setor − ISRA_Ind), cujos IS têm sinais opostos.

A predominância de *spreads* envolvendo ISRA_Ind como perna longa (três dos seis selecionados) já sinaliza que a recomendação individualizada por ação gera retornos consistentemente superiores aos das demais estratégias no cenário de calibração. Destaca-se também que os *spreads* BAH − ISRA_Setor e BAH − Índice são positivos, o que indica que, no cenário de calibração, a estratégia passiva superou o ISRA_Setor e o Índice, refletindo as características particulares do mercado brasileiro no período 2018–2019.

A Tabela 31 apresenta a distribuição dos pesos de cada portfólio. O KWP concentra cerca de 71,79% do capital no *spread* ISRA_Ind − ISRA_Setor, o de maior IS no período de calibração, enquanto o RPWP distribui de forma mais equilibrada, priorizando os *spreads* de menor volatilidade.

Ressalta-se que não se constrói um portfólio baseado em ISRA para o cenário

Tabela 31 – Distribuição dos pesos dos portfólios *Long-Short* calibrados no InS-TraD.

Spread	**EWP**	**ASRWP**	**KWP**	**RPWP**
BAH − ISRA_Setor	0,1667	0,1325	0,0968	0,1757
BAH − Índice	0,1667	0,1285	0,0537	0,1005
ISRA_Setor − Índice	0,1667	0,0442	0,0155	0,0843
ISRA_Ind − BAH	0,1667	0,0852	0,0637	0,1797
ISRA_Ind − Índice	0,1667	0,1488	0,0525	0,0849
ISRA_Ind − ISRA_Setor	0,1667	0,4607	0,7179	0,3749
Total	**1,0000**	**1,0000**	**1,0000**	**1,0000**

Nota: EWP (Pesos Iguais); ASRWP (Ponderado pelo Índice de Sharpe); KWP (Critério de Kelly); RPWP (Paridade de Risco). Apenas os 6 *spreads* com Índice de Sharpe positivo no InS-TraD foram selecionados.

OutS-TraD. Isso se deve ao fato de que OutS-TraD e InS-TraD ocorrem simultaneamente no tempo: aplicar os pesos calibrados em InS-TraD ao OutS-TraD equivaleria a utilizar informações futuras para operar no passado (*look-ahead bias*). Portanto, os pesos são aplicados exclusivamente nos cenários subsequentes e desconhecidos (InS-TesD e OutS-TesD) para testar a capacidade real de generalização da estratégia.

Por fim, vale salientar que o CDI (Certificado de Depósito Interbancário) foi adicionado ao gráfico de retorno cumulativo das estratégias apenas para efeito de comparação, visto que se trata de um ativo de renda fixa que acompanha a taxa Selic e difere dos demais.

Desempenho dos portfólios no InS-TesD:

A Figura 6 e a Tabela 32 apresentam os resultados no período de teste In-Sample (InS-TesD), que compreende o primeiro semestre de 2020 (período marcado pelo colapso dos mercados financeiros globais em decorrência da pandemia de COVID-19).

O primeiro aspecto a destacar é o contexto excepcional do período: o Ibovespa (Index) registrou retorno anualizado de $-25{,}68\%$ e *drawdown* máximo de $46{,}82\%$, refletindo a queda abrupta de março de 2020. Nesse cenário adverso, tanto o ISRA_Setor quanto o ISRA_Ind demonstram capacidade de preservação de capital: ambos registram retornos anualizados positivos ($+9{,}34\%$ e $+14{,}82\%$, respectivamente) e *drawdowns* significativamente inferiores aos do Índice e do BAH (31,29% e 27,94%, contra 46,82% do Índice e 45,81% do BAH). O ISRA_Ind apresenta ainda tempo médio de recuperação (ART) de apenas 12,75 dias, contra 36 dias do ISRA_Setor, evidenciando que a granularidade individual na recomendação confere maior agilidade de resposta em cenários de estresse.

Os portfólios *long-short* amplificam essa proteção ao operar *spreads* entre as estratégias. O **KWP** destaca-se como o melhor portfólio neste cenário, com ASR de 1,98, MDD de apenas 2,47%, MSR (1,3226) e Índice Sortino de 12,71. Enquanto o mercado colapsava

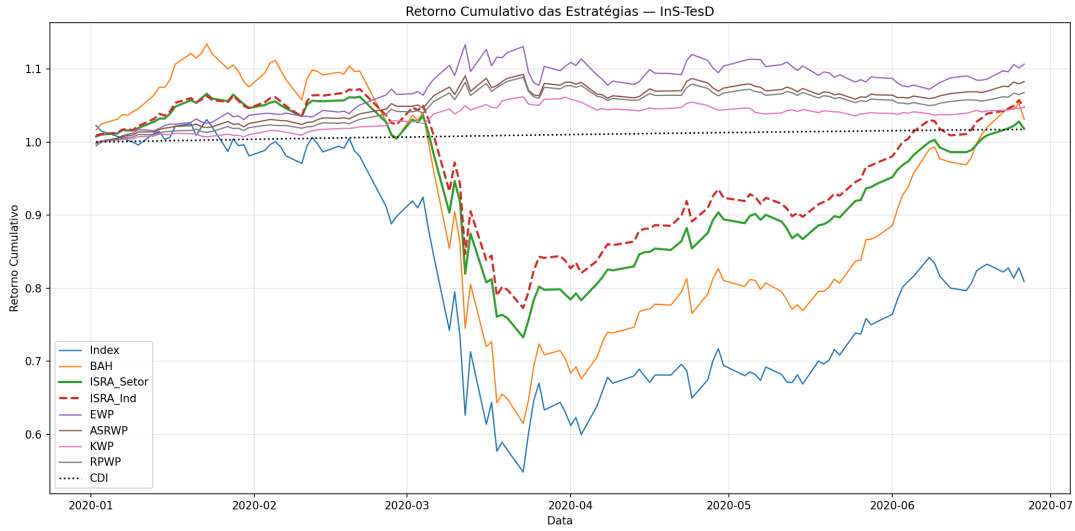

Figura 6 – Retorno acumulado das estratégias de negociação no InS-TesD (jan.–jun. 2020 — ações In-Sample).

Elaborado pelo autor.

Tabela 32 – Desempenho das estratégias de negociação no InS-TesD.

Estratégia	**WR**	**ARR**	**ASR**	**MDD**	**CAR**	**SOR**	**MSR**	**ART**
Index	0,4959	−0,2568	−0,4310	0,4682	−0,5485	−0,0251	−0,4378	39,33
BAH	0,5868	0,1902	0,3830	0,4581	0,4151	0,0193	0,3322	27,00
CDI	1,0000	0,0356	88,1653	0,0000	0,0000	0,0000	18632,4505	0,00
ISRA_Setor	0,6281	0,0934	0,2807	0,3129	0,2985	0,0133	0,2486	36,00
ISRA_Ind	0,6198	0,1482	0,4634	0,2794	0,5303	0,0223	0,3633	12,75
EWP	0,5372	**0,2186**	1,6101	0,0540	4,0476	0,1008	0,8266	12,25
ASRWP	**0,6033**	0,1677	1,9508	0,0310	**5,4053**	0,1127	0,6090	10,22
KWP	0,5620	0,0967	**1,9874**	**0,0247**	3,9099	**0,1271**	**1,3226**	9,30
RPWP	0,5702	0,1388	1,7638	0,0356	3,9036	0,1073	0,8911	**9,10**

Nota: WR: Taxa de Acerto; ARR: Retorno Anualizado; ASR: *Sharpe* Anualizado; MDD: Máximo *Drawdown*; CAR: Índice Calmar; SOR: Índice Sortino; MSR: *Sharpe* Ajustado; ART: Tempo Médio de Recuperação (dias). Valores em **negrito** indicam o melhor desempenho entre os portfólios propostos (EWP, ASRWP, KWP, RPWP). O CDI é incluído apenas como referência de baixo risco e não compete pelo destaque em negrito: seus valores de ASR e MSR decorrem da volatilidade quase nula da série e não são diretamente comparáveis em escala com as demais estratégias.

quase 47%, o KWP limitou sua queda máxima a menos de 2,47 pontos percentuais (uma redução de risco de 95% em relação ao Índice). O **EWP**, por sua vez, registrou o maior retorno anualizado entre os portfólios (ARR de 21,86%) evidenciando que a alocação igualitária entre os *spreads* captura bem o retorno absoluto disponível no cenário. O **ASRWP** também apresenta desempenho notável, com ASR de 1,95 e o segundo menor MDD (3,10%). O **RPWP** se destaca pelo menor ART (9,10 dias), indicando rápida recuperação após eventuais quedas.

É relevante notar que, neste cenário, o ISRA_Setor apresenta retorno anualizado inferior ao do BAH ($+9{,}34\%$ vs. $+19{,}02\%$), porém com MDD substancialmente menor

(31,29% vs. 45,81%), o que se reflete em melhor desempenho ajustado ao risco (CAR de 0,30 vs. 0,42 do BAH, mas ART de 36 dias vs. 27 do BAH). Isso demonstra que, durante a crise de COVID-19, a negociação ativa com o ISRA_Setor sacrificou parte do retorno em troca de menor exposição a quedas abruptas de mercado (comportamento esperado de uma estratégia baseada em sinais de *trading* que identifica janelas de não-exposição ao mercado).

Desempenho dos portfólios no OutS-TesD:

A Figura 7 e a Tabela 33 apresentam os resultados no cenário Out-of-Sample Test (OutS-TesD), que avalia a generalização das estratégias para as 49 ações fora da amostra de treinamento, no primeiro semestre de 2020.

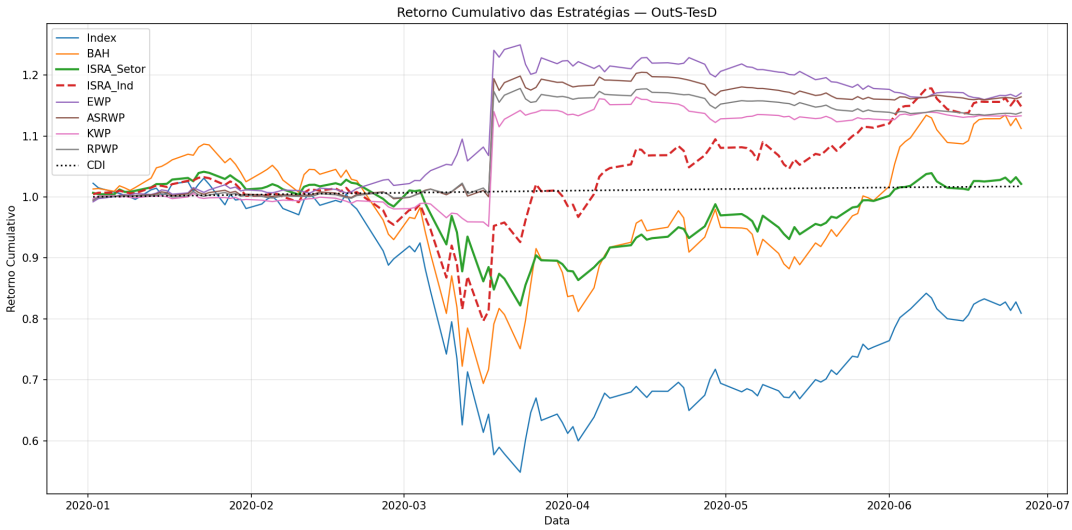

Figura 7 – Retorno acumulado das estratégias de negociação no OutS-TesD (jan.–jun. 2020 — ações Out-of-Sample).

Elaborado pelo autor.

O resultado mais expressivo do cenário OutS-TesD é o desempenho do **ISRA_Ind**: retorno anualizado de $+36{,}58\%$, ASR de 0,90 e Índice Calmar de 1,60, superando o BAH ($+36{,}01\%$; ASR 0,68) em todas as métricas ajustadas ao risco. Esse resultado é especialmente significativo: ISRA_Ind apresenta MDD de 22,84% contra 36,15% do BAH, demonstrando que a seleção individualizada de modelos reduz substancialmente o risco de queda máxima.

O ISRA_Setor, por sua vez, apresenta retorno mais modesto ($+8{,}79\%$; ASR 0,30), confirmando o padrão observado no InS-TesD: a granularidade setorial oferece proteção ao risco (MDD de 21,08%, inferior ao BAH), mas captura menos do retorno disponível. A diferença de desempenho entre ISRA_Ind e ISRA_Setor no OutS-TesD ($+36{,}58\%$ vs. $+8{,}79\%$ em ARR; 0,90 vs. 0,30 em ASR) reforça a ideia de que a recomendação individualizada por ação representa um avanço substantivo em relação à abordagem setorial.

Tabela 33 – Desempenho das estratégias de negociação no OutS-TesD.

Estratégia	WR	ARR	ASR	MDD	CAR	SOR	MSR	ART
Index	0,4959	−0,2568	−0,4310	0,4682	−0,5485	−0,0251	−0,4378	39,33
BAH	0,5372	0,3601	0,6824	0,3615	0,9963	0,0407	0,5999	17,67
CDI	1,0000	0,0356	88,1653	0,0000	0,0000	0,0000	18632,4505	0,00
ISRA_Setor	0,5289	0,0879	0,2964	0,2108	0,4171	0,0174	0,2809	22,40
ISRA_Ind	0,5455	0,3658	0,9032	0,2284	1,6014	0,0633	0,6249	11,63
EWP	**0,4876**	**0,3566**	**1,3672**	0,0719	4,9572	0,1900	−3,8622	**10,10**
ASRWP	0,4545	0,3517	1,2157	0,0377	**9,3306**	**0,2753**	−3,7400	13,38
KWP	0,4545	0,2973	1,0057	0,0544	5,4629	0,2276	**−1,5485**	15,86
RPWP	0,4793	0,2969	1,1851	**0,0373**	7,9672	0,2517	−3,2803	13,38

Nota: WR: Taxa de Acerto; ARR: Retorno Anualizado; ASR: *Sharpe* Anualizado; MDD: Máximo *Drawdown*; CAR: Índice Calmar; SOR: Índice Sortino; MSR: *Sharpe* Ajustado; ART: Tempo Médio de Recuperação (dias). Valores em **negrito** indicam o melhor desempenho entre os portfólios propostos (EWP, ASRWP, KWP, RPWP). O CDI é incluído apenas como referência de baixo risco e não compete pelo destaque em negrito: seus valores de ASR e MSR decorrem da volatilidade quase nula da série e não são diretamente comparáveis em escala com as demais estratégias.

Os portfólios *long-short* mantêm seu papel de proteção sistêmica, com MDD entre 3,73% e 7,19% (redução de 85–92% em relação ao Índice). O **EWP** se destaca pelo maior retorno anualizado entre os portfólios ($+35{,}66\%$), menor ART (10,10 dias) e pelo maior ASR (1,37). O **RPWP** apresenta o menor MDD absoluto (3,73%) e o segundo menor ART (13,38 dias), o que confirma sua característica defensiva.

Em relação ao MSR negativo observado nos portfólios *long-short* no OutS-TesD, este é um artefato da expansão de Cornish-Fisher para retornos com Índice de Sharpe muito elevado associado a distribuições com excesso de curtose. Nesses casos extremos, o termo de penalidade quadrático da fórmula

$$\text{MSR} = \text{ASR} \cdot \left[1 + \frac{S}{6}\,\text{ASR} - \frac{Ks - 3}{24}\,\text{ASR}^2\right], \tag{87}$$

torna-se dominante e inverte o sinal da métrica. Para as estratégias afetadas, o ASR tradicional é o indicador de eficiência mais fidedigno.

Em síntese, os resultados dos dois cenários de teste confirmam que: (i) a metodologia TODIM de recomendação de modelos é transferível para ações e períodos fora da amostra de calibração; (ii) a extensão para o nível individual (ISRA_Ind) representa ganho consistente em relação à abordagem setorial original (ISRA_Setor); e (iii) a combinação das recomendações algorítmicas com estratégias *long-short* oferece proteção ao risco sistêmico de forma expressiva, reduzindo o MDD em 85–95% em relação ao Índice e preservando desempenho competitivo em termos de retorno ajustado ao risco.

6 Conclusão

Este trabalho propôs e implementou um algoritmo de recomendação de modelos de negociação de ações, baseado no método TODIM, para o mercado brasileiro. O processo seguiu uma estrutura típica de Decisão Multicritério (MCDM): (i) classificação dos ativos em oito setores da B3; (ii) geração de sinais de negociação por doze modelos candidatos, incluindo algoritmos tradicionais de aprendizado de máquina e modelos neurais/de aprendizado profundo, via *Walk-Forward Analysis* (WFA), com e sem seleção de atributos; (iii) *backtesting* com oito indicadores de desempenho; (iv) recomendação do modelo ótimo por setor via TODIM; e (v) avaliação da capacidade de generalização e construção de portfólios *long-short* baseados nas recomendações. Como contribuição adicional, estendeu-se o escopo da recomendação do nível setorial ao nível de ação individual, por meio do ISRA_Ind.

A análise do ASR no período de treino (InS-TraD, 2018–2019) revelou que o Ibovespa supera os algoritmos de ML na maioria dos setores, resultado que sugere maior dificuldade dos modelos em superar o índice de mercado nesse intervalo específico. A exceção ocorreu no setor financeiro (FIN), onde o RF obteve ASR de 1,31, superando o Índice (1,30) e o BAH (1,12). No período de teste (InS-TesD, jan.–jun. 2020), marcado pelo colapso dos mercados em decorrência da pandemia de COVID-19, o cenário se inverteu: com o Ibovespa registrando ASR negativo em todos os oito setores, os melhores algoritmos superaram ambos os *benchmarks* em todos os setores. Esse resultado indica que estratégias de *trading* baseadas em ML podem agregar valor em regimes de alta volatilidade e de tendência de baixa, especialmente ao reduzir a exposição em momentos de estresse no mercado.

No cenário de calibração InS-TraD, o Ibovespa supera a maioria dos algoritmos de ML em quase todos os setores, com exceção do setor FIN, no qual o RF atinge ASR de 1,3148 contra 1,2991 do índice. Esse padrão se aproxima do observado por Lv et al. (2023) para o mercado americano, em que o S&P 500 também domina a maior parte dos setores, e diverge do observado para o mercado chinês, em que os algoritmos de ML superam o índice e o BAH na maioria das indústrias, sinalizando menor eficiência daquele mercado. Isso sugere que a B3, ao menos no período calmo pré pandemia, se comporta de forma mais próxima à hipótese de eficiência do mercado americano do que à do mercado chinês estudado no artigo original. Já no cenário de teste InS-TesD, marcado pelo colapso da pandemia de COVID-19, o Ibovespa registra ASR negativo nos oito setores, um resultado mais extremo do que qualquer um dos observados por Lv et al. (2023) para S&P 500 ou CSI 300, e em todos os setores o melhor modelo de ML supera folgadamente os dois *benchmarks*, reforçando de forma ainda mais contundente a conclusão geral do artigo de que sempre existe um algoritmo de negociação com valor de aplicação superior à estratégia passiva.

Outra diferença relevante aparece na composição dos modelos vencedores: enquanto Lv et al. (2023). relatam ASR geralmente superior dos modelos de aprendizado profundo no SPICS, a dissertação constata a predominância de modelos clássicos nas recomendações por ação individual, o que indica que, no mercado brasileiro, a vantagem de modelos mais simples sobre arquiteturas profundas é mais acentuada do que a relatada no estudo de referência.

A aplicação do TODIM às oito métricas produziu recomendações distintas por setor e cenário, confirmando a premissa fundamental do trabalho: não existe um modelo de ML universalmente superior para todos os segmentos do mercado brasileiro. Dos 32 slots de recomendação top-1, SVM e CART lideraram com cinco aparições cada, seguidos por RF, NB e GRU, com quatro aparições cada, e por LR, RNN e XGB, com três aparições cada, sem dominância consistente de nenhum algoritmo. Dois setores apresentaram maior estabilidade intrasetorial: CC, com LR recomendado em dois cenários consecutivos (InS-TesD e OutS-TraD), e UTI, com GRU como modelo ótimo tanto em InS-TraD quanto em OutS-TesD. Ao nível de ação individual, modelos clássicos de ML, incluindo o XGB como modelo de *ensemble*, dominaram 72,6% das 190 ações In-Sample e 67,3% das 49 ações Out-of-Sample, sugerindo que a complexidade adicional das redes profundas não se traduz, de forma sistemática, em ganho de desempenho no mercado brasileiro para o horizonte e a granularidade avaliados.

Além da análise do impacto da seleção de atributos, a principal contribuição deste trabalho em relação ao artigo de referência foi a extensão da recomendação à granularidade individual por ação, materializada no ISRA_Ind. Os resultados validam essa extensão de forma expressiva. No InS-TesD, o ISRA_Ind apresentou tempo médio de recuperação (ART) de 12,75 dias, em comparação aos 36 dias do ISRA_Setor, evidenciando maior agilidade de resposta em cenários de estresse. No OutS-TesD, a diferença foi ainda mais pronunciada: o ISRA_Ind obteve retorno anualizado de $+36{,}58\%$ e ASR de 0,90, superando tanto o BAH ($+36{,}01\%$; ASR 0,68) quanto o ISRA_Setor ($+8{,}79\%$; ASR 0,30) em todas as métricas ajustadas ao risco. A redução do MDD para 22,84% (contra 36,15% do BAH e 46,82% do Índice) confirma que a granularidade individual na recomendação gera benefícios concretos de preservação de capital mesmo para ativos fora da amostra de calibração.

As métricas de generalização calculadas para os três cenários de validação (InS-TesD, OutS-TraD e OutS-TesD) confirmam a robustez da estrutura proposta. Em $k = 5$, os cenários InS-TesD e OutS-TesD convergiram para valores idênticos de PR, RR e F1 (0,9091) e de CR (0,9167), indicando alta sobreposição entre os conjuntos de modelos recomendados em diferentes contextos. O MRR variou entre 0,2418 e 0,2939, e o MAP entre 0,5332 e 0,6091, faixas alinhadas às reportadas por Lv et al. (2023) para os mercados americano e chinês, o que valida a aplicabilidade da metodologia ao mercado brasileiro. Destaca-se que o cenário OutS-TraD apresentou RR@3 = 1,000, indicando que todos

os modelos presentes entre as recomendações top-3 dos setores no conjunto Out-of-Sample também aparecem no conjunto de referência InS-TraD. Esse resultado sugere forte coerência entre as preferências setoriais dos dois grupos de ativos no mesmo intervalo temporal.

Dos 12 *spreads* possíveis entre as quatro estratégias disponíveis (Índice, BAH, ISRA_Setor e ISRA_Ind), seis apresentaram Índice de Sharpe positivo no período de calibração e foram utilizados na composição dos portfólios. A predominância de *spreads* com ISRA_Ind na perna longa (três dos seis selecionados) indica uma vantagem relativa da recomendação individualizada no período de calibração. Nos cenários de teste, a combinação de recomendações algorítmicas com alocação estruturada via *long-short* demonstrou capacidade expressiva de proteção contra o risco sistêmico. No InS-TesD, enquanto o Ibovespa registrava MDD de 46,82%, o KWP limitou sua queda máxima a 2,47% (redução de 95%) com ASR de 1,99. No OutS-TesD, o EWP obteve retorno anualizado de $+35{,}66\%$ e Índice Calmar de 4,95, e o RPWP apresentou o menor MDD absoluto (3,73%), o que representa uma redução de 92% em relação ao EWP. Em ambos os cenários, os quatro portfólios *long-short* superaram simultaneamente o Índice e o BAH em ASR, validando a hipótese de que a combinação *long-short* entre estratégias algorítmicas e passivas pode gerar valor com menor exposição direcional ao mercado, especialmente durante eventos extremos de crise sistêmica.

Ademais, no plano dos portfólios *long-short*, o melhor esquema de alocação da dissertação (KWP) atinge ASR de 1,99 e redução de *drawdown* de aproximadamente 95% em relação ao Ibovespa no período de crise, superando o desempenho ajustado ao risco relatado Lv et al. (2023) para o S&P 500 (RPWP com ASR entre 0,71 e 0,93) e ficando em linha com o melhor resultado do CSI 300 (RPWP com ASR de 1,84). Um ponto de divergência importante ocorre no cenário fora da amostra em mercado de recuperação: no artigo original, os portfólios *long-short* do CSI 300 falham no OutS-TesD, apresentando retornos negativos em um mercado de tendência de alta, o que os autores atribuem à dificuldade dessas estratégias em capturar excesso de retorno em cenários ascendentes; na dissertação, por outro lado, os portfólios *long-short* permanecem lucrativos e defensivos em ambos os cenários de teste, com o EWP atingindo ARR de +35,66% no OutS-TesD, sem o mesmo colapso relatado para o mercado chinês. Por fim, a dissertação identifica um artefato metodológico não discutido no artigo original: a ocorrência de MSR negativo nos portfólios *long-short* do OutS-TesD, decorrente da instabilidade da expansão de *Cornish-Fisher* em cenários de Índice de *Sharpe* muito elevado, o que reforça a recomendação de utilizar o ASR tradicional como indicador de eficiência mais confiável nesses casos extremos.

Por fim, vale salientar que os resultados do período de teste devem ser interpretados como evidência do comportamento das estratégias sob condições de ruptura de mercado

e não como estimativa de desempenho médio em condições normais. A superação dos *benchmarks* no primeiro semestre de 2020 é condicionada ao regime de alta volatilidade e à tendência de baixa observados nesse período.

6.1 Trabalhos Futuros

Os resultados obtidos nesta dissertação abrem um conjunto de direções naturais para investigações futuras, organizadas em dois eixos complementares: extensões metodológicas imediatas e uma agenda mais ampla de pesquisa sobre o tema.

Em primeiro lugar, os experimentos realizados não incorporaram custos de transação que, no mercado brasileiro, especialmente para estratégias de *trading* diário em ações de menor liquidez, podem impactar significativamente os retornos líquidos. Trabalhos futuros devem modelar explicitamente os custos de corretagem e de derrapagem (*slippage*), tornando as estratégias avaliadas mais aderentes à realidade operacional. Em segundo lugar, o período de teste utilizado (jan.–jun. 2020) coincide com um evento de ruptura excepcional; a validação dos modelos em múltiplos ciclos de mercado, incluindo períodos de alta prolongada, baixa volatilidade e transições de regime, fortaleceria as conclusões sobre a robustez das recomendações. Em terceiro lugar, cabe destacar, como limitação deste trabalho, a presença de viés de sobrevivência (*survivorship bias*) na composição da amostra utilizada. O conjunto B3ICS foi construído a partir da lista de ações atualmente negociadas na B3, o que implica a exclusão sistemática de empresas que, ao longo do período analisado (2016 a 2020), tiveram sua listagem cancelada, entraram em processo de recuperação judicial ou faliram. Como consequência, o universo de ativos investigado tende a superestimar o desempenho médio das estratégias de negociação propostas, uma vez que companhias com trajetórias financeiras adversas, cujos retornos historicamente tendem a ser mais negativos, não estão representadas na amostra. A correção completa dessa limitação demandaria a reconstrução da base de dados a partir de séries históricas que preservem o universo efetivo de ações negociadas em cada período, como os arquivos de cotações históricas da própria B3. Por extrapolar o escopo do presente trabalho, tal reconstrução é indicada como direção relevante para pesquisas futuras. Por fim, os pesos de importância das métricas no TODIM foram definidos com base na literatura de referência e em uma ordenação fixa de preferências entre os critérios. Trabalhos futuros devem realizar uma análise de sensibilidade desses pesos, avaliando configurações alternativas que representem diferentes perfis de investidor, como os mais conservadores, os balanceados ou os agressivos.

Como desdobramento central desta pesquisa, propõe-se a investigação de Sistemas Fuzzy Evolutivos (do inglês *Evolving Fuzzy Systems*, EFS) aplicados à previsão de ativos financeiros, com foco no mercado brasileiro no cenário pós-pandemia. A principal

limitação dos doze modelos avaliados nesta dissertação é que, embora sejam reestimados periodicamente por meio da WFA, sua estrutura e seus hiperparâmetros permanecem essencialmente fixos. Assim, a adaptação a mudanças estruturais do mercado ocorre apenas de forma indireta, por meio de retreinamento em janelas, o que as torna vulneráveis ao fenômeno de deriva de dados (*data drift*) (DIXON; HALPERIN; BILOKON, 2021). Os EFS emergem como alternativa teoricamente mais adequada, pois adaptam continuamente sua estrutura e seus parâmetros à medida que novos dados chegam, reduzindo a necessidade de retreinamentos completos e permitindo uma resposta mais rápida a mudanças de regime (LUGHOFER, 2011).

Os modelos EFS a serem investigados incluem o eGNN (*Evolving Granular Neural Network*) (LEITE; COSTA; GOMIDE, 2010), o eMG (*Multivariable Gaussian Evolving Fuzzy Modeling System*) (LEITE; COSTA; GOMIDE, 2010; LEITE; COSTA; GOMIDE, 2013) e o eOGS (*Evolving Optimal Granular System*) (LEMOS; CAMINHAS; GOMIDE, 2011). Também se propõe a avaliação de comitês desses modelos, uma abordagem utilizada com resultados promissores por Amaral (2021) na previsão de preços do Bitcoin. Além disso, recomenda-se incorporar novos modelos EFS identificados ao longo da revisão da literatura sobre o tema. O desempenho dos EFS deve ser comparado diretamente ao dos algoritmos estáticos avaliados nesta dissertação.

Adicionalmente, propõe-se a extensão da metodologia de recomendação, substituindo o TODIM clássico pelo **Fuzzy-TODIM** (KROHLING; SOUZA, 2012). Enquanto o TODIM clássico opera com valores *crisp*, assumindo que os indicadores de desempenho dos modelos são conhecidos com exatidão, os indicadores financeiros estão sujeitos a múltiplas formas de incerteza simultâneas: incerteza aleatória, decorrente da variabilidade intrínseca dos retornos; incerteza epistêmica, associada à estimação a partir de amostras históricas limitadas; e incerteza subjetiva, associada à definição dos pesos dos critérios e à preferência do decisor no processo multicritério (SAHOO; CHOUDHURY; DHAL, 2026). A extensão do Fuzzy-TODIM permite modelar explicitamente essa imprecisão por meio de números triangulares fuzzy, oferecendo uma representação mais realista do processo de avaliação multicritério. Outros métodos MCDM fuzzy presentes na literatura, como Fuzzy-VIKOR, Fuzzy-TOPSIS e Fuzzy-ELECTRE, devem ser implementados para uma comparação sistemática (SAHOO; CHOUDHURY; DHAL, 2026).

Em síntese, a agenda futura proposta consiste na aplicação de EFS em conjunto com métodos MCDM Fuzzy ao problema de recomendação de modelos de *trading*, em granularidade setorial e individual por ação no mercado brasileiro pós-pandemia, com possível extensão comparativa a outros mercados, como o mercado norte-americano (S&P500).

Referências

ABOLMAKAREM, S. et al. A multi-stage machine learning approach for stock price prediction: Engineered and derivative indices. **Intelligent Systems with Applications**, v. 24, 2024. Acesso em: 16 jan. 2025. Disponível em: <https://doi.org/10.1016/j.iswa.2024.200449>. Citado na página 124.

ACHELIS, S. B. **Technical Analysis from A to Z**. 2nd. ed. [S.l.]: McGraw-Hill Education, 2000. Citado 2 vezes nas páginas 20 e 21.

ADEBIYI, A. A.; ADEWUMI, A. O.; AYO, C. K. Stock price prediction using the arima model. In: **2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation**. IEEE, 2014. p. 106–112. Accessed: 2025-06-08. Disponível em: <https://doi.org/10.1109/UKSim.2014.67>. Citado 2 vezes nas páginas 31 e 132.

AGARWAL, S. et al. Time-series forecasting using svmd-lstm: A hybrid approach for stock market prediction. **Journal of Probability and Statistics**, v. 2025, n. 1, p. 9464938, 2025. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1155/jpas/9464938>. Citado 2 vezes nas páginas 34 e 122.

AL-KHASAWNEH, M. A. et al. Stock market trend prediction using deep learning approach. **Computational Economics**, 2024. Acesso em: 15 jan. 2025. Citado na página 123.

ALALI, F.; TOLGA, A. C. Portfolio allocation with the TODIM method. **Expert Systems with Applications**, v. 124, p. 341–348, 2019. Citado na página 15.

ALJINOVIć, Z.; MARASOVIć, B.; SESTANOVIć, T. Cryptocurrency Portfolio Selection: a multicriteria approach. **Mathematics**, v. 9, n. 14, p. 1677, jul 2021. ISSN 2227-7390. Citado na página 14.

AMARAL, V. L. d. **Sistemas Fuzzy Evolutivos na Previsão e Recomendacão de Investimentos em Criptomoedas**. Dissertação (Dissertação (Mestrado)) — Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, Brazil, October 2021. Dissertação (Mestrado) apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional, CEFET-MG; Accessed: 2025-12-29. Citado na página 94.

ARMAN, A.; ARMAN, H.; HADI-VENCHEH, A. The Homogeneous MADM Methods: Is Trade-Off between Attributes Important? **Computational Intelligence and Neuroscience**, v. 2022, p. 1–11, 2022. Citado na página 14.

ATTIGERI, G. V. et al. Stock market prediction: A big data approach. In: **2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)**. Madurai, Índia: [s.n.], 2015. p. 1–5. Accessed: 2025-01-17. Disponível em: <https://doi.org/10.1109/ICCIC.2015.7435666>. Citado 2 vezes nas páginas 33 e 131.

B3. **Empresas Listadas**. 2026. Disponível em: <https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/empresas-listadas.htm>. Citado na página 51.

B3 – Brasil, Bolsa, Balcão. **Histórico de pessoas físicas na Bolsa de Valores**. B3, 2023. Published: 2023; Accessed: 2025-12-29. Disponível em: <https://www.b3.com.br>. Citado na página 3.

BARAK, S.; MODARRES, M. Developing an approach to evaluate stocks by forecasting effective features with data mining methods. **Expert Systems with Applications**, v. 42, n. 3, p. 1325–1339, 2015. Accessed: 2025-01-15. Disponível em: <https://doi.org/10.1016/j.eswa.2014.09.026>. Citado 2 vezes nas páginas 33 e 131.

BHANJA, S.; DAS, A. A black swan event-based hybrid model for indian stock markets' trends prediction. **Innovations in Systems and Software Engineering**, v. 20, p. 121–135, 2024. Disponível em: <https://doi.org/10.1007/s11334-021-00428-0>. Citado na página 125.

BIESIADA, J.; DUCH, W. Feature selection for high-dimensional data: A pearson redundancy based filter. In: **Computer Recognition Systems 2**. Berlin, Heidelberg: Springer, 2007, (Advances in Soft Computing, v. 45). p. 242–249. Citado na página 57.

BOTUNAC, I.; BOSNA, J.; MATETIć, M. Optimization of traditional stock market strategies using the lstm hybrid approach. **Information**, v. 15, p. 136, 2024. Acesso em: 16 jan. 2025. Disponível em: <https://doi.org/10.3390/info15030136>. Citado na página 125.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 4. ed. [S.l.]: John Wiley & Sons, 2013. ISBN 9781118675021. Citado na página 59.

CHANDRA, R.; HE, Y. Bayesian neural networks for stock price forecasting before and during covid-19 pandemic. **PLOS ONE**, v. 16, n. 7, p. e0253217, 2021. Accessed: 2025-01-15. Disponível em: <https://doi.org/10.1371/journal.pone.0253217>. Citado na página 129.

CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16)**. New York, NY, USA: Association for Computing Machinery, 2016. p. 785–794. KDD'16, San Francisco, CA, August 13–17, 2016. Disponível em: <https://doi.org/10.1145/2939672.2939785>. Citado 4 vezes nas páginas 43, 47, 108 e 109.

CHHAJER, P.; SHAH, M.; KSHIRSAGAR, A. The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. **Decision Analytics Journal**, Elsevier, v. 2, p. 100015, 2022. Citado 2 vezes nas páginas 34 e 128.

CRAINICH, D.; EECKHOUDT, L.; MENEGATTI, M. Some implications of common consequences in lotteries. **Journal of Risk and Uncertainty**, v. 59, n. 2, p. 185–202, 2019. Disponível em: <https://doi.org/10.1007/s11166-019-09314-4>. Citado na página 9.

CURTIS, G. Modern portfolio theory and behavioral finance. **The Journal of Wealth Management**, v. 7, n. 2, p. 16–22, jul 2004. ISSN 1534-7524, 2374-1368. Referência fornecida não inclui DOI nem URL. Citado 2 vezes nas páginas 10 e 13.

DARYL et al. Predicting stock market prices using time series sarima. In: **2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)**. IEEE, 2021. p. 302–309. Accessed: 2025-06-08. Disponível em: <https://doi.org/10.1109/ICCSAI53272.2021.9609720>. Citado 2 vezes nas páginas 32 e 132.

DAS, J. D.; OTHERS. Encoder–decoder based lstm and gru architectures for stocks and cryptocurrency prediction. **Journal of Risk and Financial Management**, v. 17, n. 5, p. 200, 2024. Citado na página 126.

DIXON, M. F.; HALPERIN, I.; BILOKON, P. **Machine Learning in Finance: From Theory to Practice**. 1. ed. Cham, Switzerland: Springer Nature Switzerland, 2021. ISBN 978-3-030-41070-4. Citado na página 94.

DOLAN, J. G. Multi-criteria clinical decision support: a primer on the use of multiple-criteria decision-making methods to promote evidence-based, patient-centered healthcare. **The Patient: Patient-Centered Outcomes Research**, v. 3, n. 4, p. 229–248, dec 2010. ISSN 1178-1653. Citado 2 vezes nas páginas 14 e 15.

DUTTA, A.; BANDOPADHYAY, G.; SENGUPTA, S. Prediction of stock performance in the indian stock market using logistic regression. **International Journal of Business and Information**, v. 7, n. 1, p. 105–136, June 2012. Accessed: 2025-06-08. Disponível em: <http://dx.doi.org/10.6702/ijbi.2012.7.1.5>. Citado 2 vezes nas páginas 32 e 132.

ENGLE, R. F.; FOCARDI, S. M.; FABOZZI, F. J. Arch/garch models in applied financial econometrics. In: LEE, C.-F.; LEE, J. C. (Ed.). **Handbook of Financial Econometrics and Statistics**. [S.l.]: Springer, 2008. cap. 60, p. 1729–1748. Accessed: 2025-06-08. Citado 2 vezes nas páginas 31 e 133.

FARAHANI, M. S.; HAJIAGHA, S. H. R. Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models. **Soft Computing**, v. 25, n. 14, p. 8483–8513, 2021. Accessed: 2025-01-14. Citado na página 129.

FARIA, E. de et al. Predicting the brazilian stock market through neural networks and adaptive exponential smoothing methods. **Expert Systems with Applications**, v. 36, n. 10, p. 12506–12509, 2009. Accessed: 2025-06-08. Disponível em: <https://doi.org/10.1016/j.eswa.2009.04.032>. Citado 2 vezes nas páginas 31 e 133.

GE, Q. Enhancing stock market forecasting: A hybrid model for accurate prediction of S&P 500 and CSI 300 future prices. **Expert Systems With Applications**, v. 260, p. 125380, 2025. Acesso em: 14 jan. 2025. Citado 2 vezes nas páginas 34 e 122.

GOMES, L. F. A. M.; LIMA, M. Todim: Basic and application to multicriteria ranking of projects with environmental impacts. **Foundations of Computing and Decision Sciences**, Poznań University of Technology, v. 16, n. 3, p. 113–127, 1991. Published: 1991. Citado 3 vezes nas páginas 2, 17 e 18.

GOMES, L. F. A. M.; LIMA, M. M. P. P. From modelling individual preferences to multicriteria ranking of discrete alternatives: a look at prospect theory and the additive difference model. **Foundations of Computing and Decision Sciences**, v. 17, n. 3, p. 171–184, 1992. Citado 2 vezes nas páginas 2 e 16.

GONG, H. et al. A new filter feature selection algorithm for classification task by ensembling pearson correlation coefficient and mutual information. **Engineering Applications of Artificial Intelligence**, v. 131, p. 107865, 2024. Citado na página 57.

GONZALO, J.; PITARAKIS, J. Estimation and model selection based inference in single and multiple threshold models. **Journal of Econometrics**, Elsevier, v. 110, n. 2, p. 319–352,

2002. Published: October 2002; Accessed: 2025-12-29. Disponível em: <https://doi.org/10.1016/S0304-4076(02)00098-2>. Citado na página 3.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, MA: MIT Press, 2016. Published: 2016. Citado 5 vezes nas páginas 43, 117, 118, 119 e 120.

GUNDUZ, H.; CATALTEPE, Z. Borsa istanbul (bist) daily prediction using financial news and balanced feature selection. **Expert Systems with Applications**, v. 42, n. 22, p. 9001–9011, 2015. ISSN 0957-4174. Citado na página 58.

GUPTA, H.; JAISWAL, A. A study on stock forecasting using deep learning and statistical models. **ArXiv**, abs/2402.06689, 2024. Disponível em: <https://api.semanticscholar.org/CorpusID:267626928>. Citado 2 vezes nas páginas 34 e 126.

HINKLE, D. E.; WIERSMA, W.; JURS, S. G. **Applied Statistics for the Behavioral Sciences**. 5. ed. Boston: Houghton Mifflin, 2003. Citado na página 57.

HINTON, G. E.; OSINDERO, S.; TEH, Y. A fast learning algorithm for deep belief nets. **Neural Computation**, MIT Press, v. 18, n. 7, p. 1527–1554, 2006. Published: July 2006. Disponível em: <https://doi.org/10.1162/neco.2006.18.7.1527>. Citado 2 vezes nas páginas 116 e 117.

HIRANSHA, M. et al. Nse stock market prediction using deep-learning models. **Procedia Computer Science**, v. 132, p. 1351–1362, 2018. Accessed: 2025-01-16. Disponível em: <https://doi.org/10.1016/j.procs.2018.05.050>. Citado 2 vezes nas páginas 33 e 130.

HOU, X. et al. An enriched time-series forecasting framework for long-short portfolio strategy. **IEEE Access**, IEEE, v. 8, p. 31992–32002, 2020. Published: 2020. Disponível em: <https://doi.org/10.1109/ACCESS.2020.2973037>. Citado na página 84.

HUANG, Y.; CAPRETZ, L. F.; HO, D. Machine learning for stock prediction based on fundamental analysis. In: **IEEE Symposium Series on Computational Intelligence (SSCI)**. [S.l.: s.n.], 2022. Acesso em: 14 jan. 2025. Citado na página 128.

IDREES, S. M.; ALAM, M. A.; AGARWAL, P. A prediction approach for stock market volatility based on time series data. **IEEE Access**, v. 7, p. 17287–17298, 2019. Accessed: 2025-01-16. Disponível em: <https://doi.org/10.1109/ACCESS.2019.2895252>. Citado 2 vezes nas páginas 31 e 132.

IEVSIEIEVA, O. et al. Financial modeling and forecasting in corporate finance management. **Economic Affairs**, v. 69, n. 1, p. 629–646, March 2024. Accessed: 2025-06-08. Disponível em: <https://doi.org/10.46852/0424-2513.2.2024.23>. Citado 2 vezes nas páginas 32 e 131.

INCE, H.; TRAFALIS, T. B. A hybrid forecasting model for stock market prediction. **Economic Computation and Economic Cybernetics Studies and Research**, v. 51, n. 3, p. 263–280, 2017. Citado 2 vezes nas páginas 33 e 131.

Instituto Brasileiro de Geografia e Estatística. **Relação da população dos municípios para publicação no DOU em 2023**. IBGE, 2023. Dados populacionais oficiais; Accessed: 2025-12-29. Disponível em: <https://ftp.ibge.gov.br/Informacoes_Gerais_e_Referencia/Relacao_da_Populacao_dos_Municipios_para_publicacao_no_DOU_em_2023/POP_TCU_2023_Brasil_e_UFs_POP2022_Malha2023.pdf>. Citado na página 4.

IZZELDIN, M. et al. The impact of the russian-ukrainian war on global financial markets. **International Review of Financial Analysis**, Elsevier, v. 87, p. 102598, 2023. Available online: 2023; Accessed: 2025-12-29. Disponível em: <https://doi.org/10.1016/j.irfa.2023.102598>. Citado na página 3.

JAIN, V.; SAHAY, R. R.; NUPUR. Integrated robust em and todim approach with teaching learning based portfolio optimization. **Computational Economics**, Springer Science+Business Media, LLC, 2025. Accepted: 31 May 2025; Accessed: 2025-09-08. Disponível em: <https://doi.org/10.1007/s10614-025-11013-z>. Citado 2 vezes nas páginas 36 e 134.

KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. **Econometrica**, v. 47, n. 2, p. 263–291, 1979. Accessed: 2025-08-23. Disponível em: <https://doi.org/10.2307/1914185>. Citado 3 vezes nas páginas 7, 10 e 12.

KANIEL, R.; SAAR, G.; TITMAN, S. Individual investor trading and stock returns. **The Journal of Finance**, Wiley for the American Finance Association, v. 63, n. 1, p. 273–310, 2008. Published: February 2008; Accessed: 2025-12-29. Disponível em: <https://doi.org/10.1111/j.1540-6261.2008.01333.x>. Citado na página 3.

KORCZAK, J.; HERNES, M. Deep learning for financial time series forecasting in a-trader system. In: **Proceedings of the 2017 Federated Conference on Computer Science and Information Systems (FedCSIS)**. Polish Information Processing Society / ACSIS, 2017. (Annals of Computer Science and Information Systems, v. 11), p. 905–912. FedCSIS 2017, Prague, Czech Republic, Sept 3–6 2017; Accessed: 2025-12-29. Disponível em: <http://dx.doi.org/10.15439/2017F449>. Citado na página 3.

KOU, G. et al. Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. **Applied Soft Computing**, v. 86, p. 105836, 2020. Citado na página 15.

KOU, G.; PENG, Y.; WANG, G. Evaluation of clustering algorithms for financial risk analysis using MCDM methods. **Information Sciences**, v. 275, p. 1–12, 2014. Citado na página 15.

KROHLING, R. A.; SOUZA, T. T. M. de. Combining prospect theory and fuzzy numbers to multi-criteria decision making. **Expert Systems with Applications**, Elsevier, v. 39, p. 11487–11493, 2012. Published: 2012. Disponível em: <https://doi.org/10.1016/j.eswa.2012.04.007>. Citado na página 94.

LANIADO, D. et al. Gender homophily in online dyadic and triadic relationships. **EPJ Data Science**, SpringerOpen, v. 5, p. 19, 2016. Article number 19; Published: 2016. Disponível em: <https://doi.org/10.1140/epjds/s13688-016-0083-3>. Citado na página 43.

LEE, K.; YOO, S.; JIN, J. J. Neural network model vs. sarima model in forecasting korean stock price index (kospi). **Issues in Information Systems**, v. 8, n. 2, p. 372–378, 2007. Accessed: 2025-06-08. Disponível em: <https://doi.org/10.48009/2_iis_2007_372-378>. Citado 2 vezes nas páginas 30 e 133.

LEITE, D.; COSTA, P.; GOMIDE, F. Evolving granular neural network for semi-supervised data stream classification. In: **Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN 2010)**. Barcelona, Spain: Institute of Electrical and Electronics Engineers (IEEE), 2010. p. 1–8. IJCNN 2010. Citado na página 94.

LEITE, D.; COSTA, P.; GOMIDE, F. Evolving granular neural networks from fuzzy data streams. **Neural Networks**, Elsevier, v. 38, p. 1–16, 2013. Published: 2013. Disponível em: <https://doi.org/10.1016/j.neunet.2012.11.007>. Citado na página 94.

LEMOS, A.; CAMINHAS, W.; GOMIDE, F. Multivariable gaussian evolving fuzzy modeling system. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 19, n. 1, p. 91–104, 2011. Published: 2011. Disponível em: <https://doi.org/10.1109/TFUZZ.2010.2087381>. Citado na página 94.

LEONETI, A. B.; GOMES, L. F. A. M. A novel version of the TODIM method based on the exponential model of prospect theory: The ExpTODIM method. **European Journal of Operational Research**, v. 295, n. 3, p. 1042–1055, 2021. Citado 2 vezes nas páginas 15 e 16.

LUGHOFER, E. **Evolving Fuzzy Systems: Methodologies, Advanced Concepts and Applications**. Berlin, Heidelberg: Springer, 2011. ISBN 978-3-642-18086-6. Citado na página 94.

LV, D. et al. Recommendation algorithm of industry stock trading model with todim. **International Journal of Information Technology & Decision Making**, World Scientific Publishing, v. 23, n. 3, p. 1301–1334, 2023. Accessed: 2025-09-08. Disponível em: <https://doi.org/10.1142/S0219622023500402>. Citado 19 vezes nas páginas 1, 3, 5, 15, 16, 36, 38, 46, 50, 61, 63, 77, 78, 79, 83, 90, 91, 92 e 134.

LV, D. et al. Selection of the optimal trading model for stock investment in different industries. **PLoS ONE**, Public Library of Science, v. 14, n. 2, p. 1–20, 2019. Published: February 13, 2019; Accessed: 2025-12-29. Disponível em: <https://doi.org/10.1371/journal.pone.0212137>. Citado na página 1.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to Information Retrieval**. Cambridge: Cambridge University Press, 2008. ISBN 9780521865715. Citado na página 63.

MARQUéS, A. I.; GARCIA, V.; SáNCHEZ, J. S. Ranking-based MCDM models in financial management applications: analysis and emerging challenges. **Progress in Artificial Intelligence**, v. 9, n. 3, p. 171–193, sep 2020. ISSN 2192-6352, 2192-6360. Citado na página 16.

MEHTA, P.; PANDYA, S.; KOTECHA, K. Harvesting social media sentiment analysis to enhance stock market prediction using deep learning. **PeerJ Computer Science**, v. 7, p. e476, 2021. Accessed: 2025-01-15. Disponível em: <https://doi.org/10.7717/peerj-cs.476>. Citado 2 vezes nas páginas 34 e 129.

MERH, N.; SAXENA, V. P.; PARDASANI, K. R. A comparison between hybrid approaches of ann and arima for indian stock trend forecasting. **Business Intelligence Journal**, v. 3, n. 2, p. 23–43, 2010. Accessed: 2025-06-08. Disponível em: <https://www.researchgate.net/publication/45602108>. Citado 2 vezes nas páginas 31 e 132.

MUKAKA, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. **Malawi Medical Journal**, v. 24, n. 3, p. 69–71, 2012. Citado na página 57.

MURPHY, K. P. **Machine Learning: A Probabilistic Perspective**. Cambridge, MA: MIT Press, 2012. Published: 2012. Citado 4 vezes nas páginas 43, 114, 115 e 116.

MUTINDA, J. K.; LANGAT, A. K. Stock price prediction using combined garch-ai models. **Scientific African**, v. 26, p. e02374, 2024. Acesso em: 16 jan. 2025. Disponível em: <https://doi.org/10.1016/j.sciaf.2024.e02374>. Citado 2 vezes nas páginas 34 e 124.

NABIPOUR, M. et al. Deep learning for stock market prediction. **Entropy**, v. 22, n. 8, p. 840, 2020. Accessed: 2025-01-14. Disponível em: <https://doi.org/10.3390/e22080840>. Citado 2 vezes nas páginas 34 e 130.

NI, J.; ZHANG, C. An efficient implementation of the backtesting of trading strategies. In: PAN, Y. et al. (Ed.). **Parallel and Distributed Processing and Applications**. Berlin, Heidelberg: Springer, 2005. (Lecture Notes in Computer Science, v. 3758), p. 126–131. Citado na página 46.

PADHI, D. K. et al. An intelligent fusion model with portfolio selection and machine learning for stock market prediction. **Computational Intelligence and Neuroscience**, v. 2022, p. 1–18, 2022. Disponível em: <https://doi.org/10.1155/2022/7588303>. Citado na página 127.

PAN, H.; TANG, Y.; WANG, G. A stock index futures price prediction approach based on the multi-garch-lstm mixed model. **Mathematics**, v. 12, n. 1677, 2024. Acesso em: 14 jan. 2025. Disponível em: <https://doi.org/10.3390/math12111677>. Citado 2 vezes nas páginas 34 e 125.

PANG, X. et al. An innovative neural network approach for stock market prediction. **The Journal of Supercomputing**, v. 76, n. 3, p. 2098–2118, 2020. Accessed: 2025-01-16. Disponível em: <https://doi.org/10.1007/s11227-017-2228-y>. Citado 2 vezes nas páginas 34 e 130.

PENG, H.; LONG, F.; DING, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 27, n. 8, p. 1226–1238, 2005. Citado na página 56.

PHUOC, T.; OTHERS. Applying machine learning algorithms to predict the stock price trend in the stock market – the case of vietnam. **Humanities & Social Sciences Communications**, v. 11, n. 1, 2024. Citado na página 126.

PINDYCK, R. S.; RUBINFELD, D. L. **Microeconomia**. 8. ed. São Paulo: Pearson Education do Brasil, 2013. Citado 2 vezes nas páginas 7 e 8.

PUPPO, B. et al. A multicriteria decision trading system based on prospect theory: A risk return analysis of the todim method. **Processes**, v. 10, p. 609, mar 2022. Received: 31 January 2022; Accepted: 22 February 2022; Published: 21 March 2022. Open access (CC BY 4.0). Disponível em: <https://doi.org/10.3390/pr10030609>. Citado 4 vezes nas páginas 11, 20, 35 e 133.

PUPPO, B. D. **Hybrid Model for Selecting Investment Assets Using the TODIM-θ Method and Modern Portfolio Theory**. Tese (Doctor of Science Thesis) — Instituto Tecnológico de Aeronáutica (ITA) and Universidade Federal de São Paulo (UNIFESP), São José dos Campos, Brazil, 2024. Graduate Program in Operations Research, Field of Management and Decision Support. Disponível em: <https://repositorio.ita.br/>. Citado 3 vezes nas páginas 12, 13 e 17.

QIAN, E. **Risk parity portfolios: Efficient portfolios through true diversification**. Boston, MA, 2005. Technical Report. Citado na página 85.

QIU, M.; SONG, Y.; AKAGI, F. Application of artificial neural network for the prediction of stock market returns: The case of the japanese stock market. **Chaos, Solitons and Fractals**, v. 85, p. 1–7, 2016. Accessed: 2025-01-16. Disponível em: <https://doi.org/10.1016/j.chaos.2016.01.004>. Citado 2 vezes nas páginas 33 e 131.

RADOVIC, M. et al. Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. **BMC Bioinformatics**, v. 18, n. 1, p. 9, 2017. Citado na página 56.

RAMOS, I. C. d. O.; FREIRE, F. H. M. d. A.; BARBOSA, D. A. **Estatística não-paramétrica utilizando o R**. 2018. Natal, julho de 2018. Citado na página 70.

SAATY, T. L. Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World. **European Journal of Operational Research**, v. 48, p. 9–26, 1990. Citado na página 17.

SAFARI, A.; BADAMCHIZADEH, M. A. Deepinvesting: Stock market predictions with a sequence-oriented bilstm stacked model – a dataset case study of amzn. **Intelligent Systems with Applications**, v. 24, p. 200439, 2024. Acesso em: 16 jan. 2025. Disponível em: <https://doi.org/10.1016/j.iswa.2024.200439>. Citado na página 124.

SAFARI, A.; GHAEMI, S. NeuroFuzzyMan: A hybrid neuro-fuzzy BiLSTM stacked ensemble model for financial forecasting and analysis: Dataset case studies on JPMorgan, AMZN and TSLA. **Expert Systems with Applications**, v. 266, p. 126037, 2025. Acesso em: 14 jan. 2025. Citado 2 vezes nas páginas 34 e 122.

SAHOO, S. K.; CHOUDHURY, B. B.; DHAL, P. R. A comprehensive review of fuzzy multiple criteria decision-making (mcdm) methods: Advancements, applications, and future directions. **Spectrum of Decision Making and Applications**, v. 00, n. 00, p. 1–15, 2026. Citado na página 94.

SAXENA, U. R.; SHARMA, P.; GUPTA, G. Comprehensive study of machine learning algorithms for stock market prediction during covid-19. **Journal of Computers, Mechanical and Management**, v. 2, n. 1, p. 1–7, 2023. Accessed: 2025-01-14. Citado 2 vezes nas páginas 34 e 127.

SCHMIDT, U. **Alternatives to Expected Utility: Formal Theories**. Springer, Boston, MA, 2004. 757–837 p. Disponível em: <https://doi.org/10.1007/978-1-4020-7964-1_2>. Citado na página 8.

SHABAN, W. M.; ASHRAF, E.; SLAMA, A. E. SMP-DL: a novel stock market prediction approach based on deep learning for effective trend forecasting. **Neural Computing and Applications**, v. 36, p. 1849–1873, 2024. Acesso em: 15 jan. 2025. Citado na página 124.

SHAHID, M. Covid-19 and adaptive behavior of returns: evidence from commodity markets. **Humanities and Social Sciences Communications**, Springer Nature, v. 9, n. 1, p. 1–11, 2022. Published: 2022; Accessed: 2025-12-29. Disponível em: <https://doi.org/10.1057/s41599-022-01106-5>. Citado na página 3.

SILVA, C. A. d. et al. Utilização do método multicritério todim-fse para classificação de base de brigada. In: **Anais do 27. Simpósio de Pesquisa Operacional e Logística da Marinha, Rio de Janeiro**. São Paulo: Edgard Blücher, 2014. p. 419–430. Anais publicados pela editora; referência original não indica DOI nem URL. Citado na página 11.

SINGH, A. et al. Application of neural network to technical analysis of stock market prediction. In: **2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)**. IEEE, 2022. p. 302–306. Disponível em: <https://doi.org/10.1109/ICIEM54221.2022.9853162>. Citado na página 128.

SINGH, J.; KHUSHI, M. Feature learning for stock price prediction shows a significant role of analyst rating. **Data**, v. 4, n. x, p. 1–20, 2021. Accessed: 2025-01-14. Disponível em: <https://doi.org/10.3390/xxxxx>. Citado na página 129.

SIRIGIRI, P.; HOTA, H. S.; SHARMA, L. K. Students Performance Evaluation using MCDM Methods through Customized Software. **International Journal of Computer Applications**, v. 130, n. 15, p. 11–14, 2015. Citado na página 14.

SOARE, C. A. P.; CHINELLI, C. K. Multicriteria decision-making methods. In: **Materials Selection for Sustainability in the Built Environment Environmental, Social and Economic Aspects**. [S.l.]: Elsevier BV, 2024. p. 299–317. ISBN 978-0-323-95122-7. Citado na página 14.

TAHERDOOST, H.; MADANCHIAN, M. Multi-Criteria Decision Making (MCDM) Methods and Concepts. **Encyclopedia**, v. 3, n. 1, p. 77–87, 2023. Citado 2 vezes nas páginas 14 e 15.

TASHAKKORI, A. et al. Enhancing stock market prediction accuracy with recurrent deep learning models: A case study on the CAC40 index. **World Journal of Advanced Research and Reviews**, v. 23, n. 1, p. 2309–2321, 2024. Acesso em: 15 jan. 2025. Citado na página 123.

THOMSETT, M. **Technical Analysis of Stock Trends Explained: An Easy-to-Understand System for Successful Trading**. [S.l.]: Ethan Hathaway Co Ltd, 2012. Copyright © 2012. Companion course available at http://www.ethanhathawayonline.com/technicalanalysisonlinetrainingcours. Citado na página 21.

THORP, E. O. The kelly criterion in blackjack, sports betting, and the stock market. In: **The Kelly Capital Growth Investment Criterion: Theory and Practice**. [S.l.]: World Scientific, 2011. p. 789–832. Published: 2011. Citado na página 84.

TVERSKY, A.; KAHNEMAN, D. Advances in prospect theory: Cumulative representation of uncertainty. **Journal of Risk and Uncertainty**, v. 5, n. 4, p. 297–323, 1992. Accessed: 2025-08-24. Disponível em: <https://doi.org/10.1007/BF00122574>. Citado 2 vezes nas páginas 13 e 52.

TVERSKY, A.; KAHNEMAN, D. Advances in prospect theory: cumulative representation of uncertainty. In: ARLó-COSTA, H.; HENDRICKS, V. F.; BENTHEM, J. van (Ed.). **Readings in Formal Epistemology: Sourcebook**. Cham: Springer International Publishing, 2016. p. 493–519. ISBN 978-3-319-20450-5, 978-3-319-20451-2. Reprint of the original article (Journal of Risk and Uncertainty, 1992). Citado na página 11.

United Nations Conference on Trade and Development. **World Investment Report 2020: International Production Beyond the Pandemic**. United Nations, 2020. Accessed: 2025-12-29. Disponível em: <https://unctad.org/publication/world-investment-report-2020>. Citado na página 3.

UçKAN, T. Integrating PCA with deep learning models for stock market forecasting: An analysis of Turkish stocks markets. **Journal of King Saud University - Computer and Information Sciences**, v. 36, 2024. Acesso em: 14 jan. 2025. Citado na página 123.

VELASQUEZ, M.; HESTER, P. T. An analysis of multi-criteria decision making methods. **International Journal of Operations Research**, v. 10, n. 2, p. 56–66, 2013. Disponível em: <http://www.orstw.org.tw/ijor/vol10no2/ijor_vol10_no2_p56_p66.pdf>. Citado 2 vezes nas páginas 14 e 15.

VULLAM, N. et al. Prediction and analysis using a hybrid model for stock market. In: **2023 3rd International Conference on Intelligent Technologies (CONIT)**. [S.l.: s.n.], 2023. p. 1–8. Accessed: 2025-01-14. Citado 2 vezes nas páginas 34 e 127.

WALLENIUS, J. et al. Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. **Management Science**, v. 54, n. 7, p. 1336–1349, jul 2008. ISSN 0025-1909, 1526-5501. Citado na página 14.

WANG, C.-N. et al. Simulation-based optimized weighting TODIM decision-making approach for national oil company global benchmarking. **IEEE Transactions on Engineering Management**, p. 1–15, 2022. Citado 3 vezes nas páginas 16, 35 e 134.

WANG, J.; WANG, X. Covid-19 and financial market efficiency: Evidence from an entropy-based analysis. **Finance Research Letters**, Elsevier, v. 42, p. 101888, 2021. Available online: 2020; Accessed: 2025-12-29. Disponível em: <https://doi.org/10.1016/j.frl.2020.101888>. Citado na página 3.

WANG, S.; LIU, J. Extension of the TODIM method to intuitionistic linguistic multiple attribute decision making. **Symmetry**, v. 9, n. 6, p. 95, jun 2017. ISSN 2073-8994. Citado na página 17.

XP Investimentos. **Relatório de Investidores Pessoa Física na Bolsa de Valores**. XP Investimentos, 2023. Published: 2023; Accessed: 2025-12-29. Disponível em: <https://conteudos.xpi.com.br>. Citado na página 3.

XP Investimentos. **Investidores na B3 chegam a 19,1 milhões; número de PF em renda fixa, FIIs e ações cresce**. XP Investimentos, 2024. XP Expert (conteúdo online). Publicado: 23 Feb 2024; Atualizado: 23 Feb 2024; Accessed: 2025-12-29. Disponível em: <https://conteudos.xpi.com.br/conteudos-gerais/b3-investidores-renda-fixa-variavel-fiis-dez-2023/>. Citado na página 4.

YANG, J.; WANG, Y.; LI, X. Prediction of stock price direction using the lasso-lstm model combines technical indicators and financial sentiment analysis. **PeerJ Computer Science**, PeerJ, v. 8, p. e1148, 2022. Citado na página 127.

YASIN, S.; PASCHKE, A.; QUNDUS, J. A. Prediction of stock prices with automated reinforced learning algorithms. **Preprint submitted to Expert Systems**, 2024. Citado 2 vezes nas páginas 34 e 126.

YENIREDDY, A.; OTHERS. Stock market index prediction based on market trend using lstm. **Indonesian Journal of Electrical Engineering and Computer Science**, v. 35, n. 3, p. 1601, 2024. Citado na página 126.

YIN, J. et al. An extended TODIM method for project manager's competency evaluation. **Journal of Civil Engineering and Management**, v. 25, n. 7, p. 673–686, 2019. Citado na página 16.

YOSHINAGA, C. E.; RAMALHO, T. B. Finâncias comportamentais no brasil: uma aplicação da teoria da perspectiva em potenciais investidores. **Review of Business Management**, v. 16, n. 53, p. 594–615, dec 2014. ISSN 1806-4892, 1983-0807. Referência fornecida não inclui DOI nem URL. Citado 2 vezes nas páginas 8 e 10.

YU, X. et al. A TODIM-based approach for the assessment of social vulnerability in Jiangsu province from 2012 to 2017. **Journal of Intelligent Fuzzy Systems**, v. 41, p. 7457–7471, 2021. Citado na página 15.

ZAICHENKO, A. et al. The battle of information representations: Comparing sentiment and semantic features for forecasting market trends. **CoRR**, abs/2303.14221, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2303.14221>. Citado 2 vezes nas páginas 34 e 127.

ZARGHAMI, M.; SZIDAROVSZKY, F. Introduction to multicriteria decision analysis. In: **Multicriteria Analysis**. [S.l.]: Springer-Verlag Berlin Heidelberg, 2011. cap. 1, p. 1–12. Citado na página 14.

ZHANG, J. The impact of prospect theory on asset forecasting: A comparative study based on random forest and linear regression. In: **Proceedings of ICFTBA 2024 Workshop: Finance in the Age of Environmental Risks and Sustainability**. [s.n.], 2024. p. 130–138. Accessed: 2025-08-24. Disponível em: <https://doi.org/10.54254/2754-1169/94/2024OX0184>. Citado 2 vezes nas páginas 35 e 133.

ZHANG, K. et al. Stock market prediction based on generative adversarial network. **Procedia Computer Science**, v. 147, p. 400–406, 2019. Accessed: 2025-01-15. Disponível em: <https://doi.org/10.1016/j.procs.2019.01.256>. Citado 2 vezes nas páginas 33 e 130.

ZHANG, Y.; LI, X.; GUO, S. Portfolio selection problems with Markowitz's mean–variance framework: a review of literature. **Fuzzy Optimization and Decision Making**, v. 17, n. 2, p. 125–158, jun 2018. ISSN 1568-4539, 1573-2908. Disponível em: <https://doi.org/10.1007/s10700-017-9266-z>. Citado na página 9.

ZHENG, J. et al. The Random Forest Model for Analyzing and Forecasting the US Stock Market under the Background of Smart Finance. In: **Proceedings of the 3rd International Academic Conference on Blockchain, Information Technology and Smart Finance (ICBIS 2024)**. [s.n.], 2024. (Atlantis Highlights in Computer Sciences), p. 83–90. Open Access under Creative Commons Attribution-NonCommercial 4.0 International License. Disponível em: <https://doi.org/10.2991/978-94-6463-419-8_11>. Citado na página 123.

ZHOU, Z. **Machine Learning**. Cham, Switzerland: Springer Nature, 2021. Published: 2021. Citado 8 vezes nas páginas 43, 107, 108, 110, 111, 112, 113 e 114.

Apêndices

APÊNDICE A – Sobre os modelos de *Machine Learning* utilizados no trabalho

A.1 Modelos de Aprendizado de Máquina

A seção inicia-se abordando os modelos tradicionais de aprendizado de máquina e, em seguida, introduz os modelos de aprendizado profundo, fundamentados em redes neurais profundas. Os modelos apresentados nesta seção foram selecionados por representarem diferentes paradigmas de aprendizado, incluindo abordagens lineares, probabilísticas, baseadas em árvores e neurais. Essa diversidade é relevante neste trabalho, pois permite avaliar como diferentes perfis de modelagem se comportam em distintos setores do mercado acionário, fornecendo um conjunto heterogêneo de alternativas para o algoritmo de recomendação baseado no método TODIM.

A.1.1 *Support Vector Machine*

O *Support Vector Machine* (SVM) é um algoritmo de aprendizado de máquina supervisionado, amplamente utilizado para tarefas de classificação. O objetivo fundamental do modelo é encontrar uma fronteira de decisão, conhecida como hiperplano separador, capaz de separar amostras de diferentes classes no espaço de dados (ZHOU, 2021).

Embora existam múltiplas formas de separar os dados, o SVM se destaca por buscar a separação mais robusta possível. A intuição por trás do modelo é escolher a fronteira que passa exatamente pelo “meio” das classes, garantindo maior tolerância a ruídos e melhor capacidade de generalização para novos dados (ZHOU, 2021).

O conceito central do SVM reside nos vetores de suporte e na margem. Os vetores de suporte são os pontos de dados (amostras de treinamento) mais próximos da fronteira de decisão. Eles são chamados assim porque, efetivamente, “suportam” ou definem a posição do hiperplano (ZHOU, 2021).

A distância entre esses vetores de suporte e o hiperplano é denominada margem. O SVM opera sob o princípio de maximização da margem, buscando posicionar a fronteira de decisão de modo que a distância seja a maior possível. Uma margem mais larga implica que o modelo é mais seguro em suas classificações e menos propenso a erros causados por pequenas perturbações nos dados (ZHOU, 2021).

Em cenários reais, os dados raramente são perfeitamente separáveis devido a ruídos ou sobreposições entre as classes. Tentar forçar uma separação perfeita (chamada de *hard*

margin) geralmente leva ao *overfitting*, em que o modelo se ajusta excessivamente aos dados de treino e falha na generalização. Para contornar isso, o SVM utiliza o conceito de margem suave. Essa abordagem permite que o modelo tolere alguns erros de classificação (permitindo que alguns pontos fiquem dentro da margem ou no lado errado da fronteira) em troca de uma fronteira de decisão globalmente mais estável. O equilíbrio entre maximizar a margem e minimizar os erros é controlado por um parâmetro de regularização, que define o quão rígido o modelo deve ser em relação aos erros (ZHOU, 2021).

Muitos problemas complexos não podem ser resolvidos com uma simples linha reta ou plano (separação linear). Para lidar com isso, o SVM emprega uma técnica chamada *kernel trick*. Essa técnica projeta os dados originais para um espaço de dimensão superior (um espaço de características). Nesse espaço mais complexo, dados que antes estavam misturados podem se tornar linearmente separáveis. Uma vez encontrada a separação nesse espaço, ela corresponde a uma fronteira não linear (curva e complexa) quando projetada de volta ao espaço original. Existem diferentes tipos de funções de *kernel* (como polinomiais ou gaussianas/RBF) que determinam como essa projeção é realizada (ZHOU, 2021).

Além da classificação, os princípios do SVM são aplicados a problemas de regressão por meio do *Support Vector Regression* (SVR). Diferentemente da regressão tradicional, que busca minimizar o erro de cada ponto, o SVR define uma “margem de tolerância” ao redor da previsão. Erros que caem nessa margem são ignorados pelo modelo, permitindo focar nas tendências principais dos dados e desconsiderar pequenas flutuações irrelevantes (ZHOU, 2021).

A.1.2 *Extreme Gradient Boosting*

O *XGBoost* é um sistema de aprendizado de máquina projetado para ser altamente escalável e eficiente, que implementa o método de *Gradient Tree Boosting*. Amplamente adotado em competições de ciência de dados e em aplicações industriais, o modelo destaca-se por fornecer resultados de estado da arte em tarefas de classificação e regressão, lidando de forma eficiente com dados esparsos e com grandes volumes de informação (CHEN; GUESTRIN, 2016).

O modelo opera por meio de uma abordagem de *ensemble*, na qual múltiplos modelos simples (neste caso, árvores de regressão) são combinados para formar um preditor robusto. A predição final para um determinado exemplo é obtida pela soma das pontuações atribuídas pelas árvores individuais (CHEN; GUESTRIN, 2016).

Diferentemente dos métodos tradicionais de *Gradient Boosting*, o *XGBoost* otimiza uma função objetivo regularizada. Essa função é composta por dois componentes principais (CHEN; GUESTRIN, 2016):

- **Função de Perda Convexa (l):** Mede a discrepância entre a predição do modelo e o valor-alvo, representando o erro de treinamento.
- **Termo de Regularização (Ω):** Penaliza a complexidade do modelo, considerando o número de folhas e a magnitude dos pesos associados a elas.

A inclusão do termo de regularização contribui para suavizar os pesos finais aprendidos, prevenindo o *overfitting* e favorecendo modelos mais simples, com maior capacidade de generalização. O treinamento do modelo ocorre de forma aditiva, em estágios sucessivos. A cada iteração, uma nova árvore é adicionada ao conjunto com o objetivo de corrigir os erros cometidos pelas árvores anteriores Chen e Guestrin (2016). Do ponto de vista matemático, o *XGBoost* utiliza uma aproximação de segunda ordem da função objetivo. Isso implica o uso tanto do gradiente (primeira derivada) quanto da Hessiana (segunda derivada) da função de perda para orientar o processo de otimização e definir a estrutura da nova árvore. Essa estratégia possibilita uma convergência mais rápida e precisa em comparação com métodos que utilizam apenas informações de primeira ordem (CHEN; GUESTRIN, 2016).

Um dos desafios centrais no treinamento de árvores de decisão consiste em identificar o melhor ponto de divisão (*split*) para cada nó. O XGBoost emprega estratégias avançadas para resolver esse problema (CHEN; GUESTRIN, 2016):

1. **Algoritmo Guloso Exato:** Para conjuntos de dados que cabem inteiramente na memória, o algoritmo avalia todos os possíveis pontos de divisão em cada característica, selecionando aquele que maximiza o ganho de informação.
2. **Algoritmo Aproximado:** Para grandes volumes de dados, são propostos candidatos a pontos de divisão com base nos percentis da distribuição das características.
3. ***Weighted Quantile Sketch*:** Método introduzido pelo *XGBoost* para o cálculo eficiente de quantis em dados ponderados, garantindo precisão teórica mesmo em ambientes distribuídos.

Dados do mundo real frequentemente apresentam valores ausentes ou são esparsos por natureza, como ocorre na codificação *one-hot*. O *XGBoost* incorpora um algoritmo nativo, ciente da esparsidade, para lidar com essas situações (CHEN; GUESTRIN, 2016).

Para cada nó da árvore, o algoritmo aprende uma direção padrão. Durante a fase de predição, caso um valor de uma característica esteja ausente em um determinado exemplo, este é automaticamente encaminhado para a direção padrão aprendida. Essa estratégia não apenas trata os dados faltantes de forma consistente, como também reduz significativamente o custo computacional, uma vez que apenas os valores não ausentes precisam ser processados (CHEN; GUESTRIN, 2016).

Além das inovações algorítmicas, o *XGBoost* incorpora otimizações de sistema relevantes (CHEN; GUESTRIN, 2016):

- **Estrutura de Blocos em Colunas:** Os dados são organizados em blocos de memória

com colunas pré-ordenadas, o que permite um processamento paralelo eficiente.

- **Computação *Out-of-Core*:** Para conjuntos de dados que excedem a capacidade da memória RAM, o sistema divide os dados em blocos armazenados em disco, utilizando compressão e *sharding* (fragmentação) para maximizar a taxa de leitura e viabilizar o treinamento em *hardware* limitado.

A.1.3 *Random Forest*

O *Random Forest* (RF) é um método de aprendizado de conjunto (*ensemble learning*) considerado uma extensão direta do algoritmo *Bagging*. Ele é amplamente reconhecido como um método representativo do estado da arte devido ao seu desempenho robusto em aplicações do mundo real, além de seu baixo custo computacional e facilidade de implementação (ZHOU, 2021).

A principal inovação do *Random Forest* em relação ao *Bagging* tradicional reside na introdução de aleatoriedade adicional na construção das árvores. Enquanto o *Bagging* gera diversidade principalmente por meio da reamostragem de dados (*bootstrap*), o *Random Forest* incorpora também a seleção aleatória de características (ZHOU, 2021).

O processo de divisão dos nós difere das árvores de decisão convencionais da seguinte maneira (ZHOU, 2021):

- **Árvores Tradicionais:** Para dividir um nó, o algoritmo avalia todas as características disponíveis no conjunto de dados para encontrar a divisão ótima.
- ***Random Forest*:** Para cada nó da árvore, o algoritmo seleciona aleatoriamente um subconjunto de k características do conjunto total. A melhor divisão é, então, escolhida apenas dentro desse subconjunto restrito de candidatos.

O parâmetro k é fundamental para controlar o grau de aleatoriedade introduzido no modelo (ZHOU, 2021):

- Se $k = d$ (onde d é o número total de características), a construção da árvore é equivalente à de uma árvore de decisão tradicional utilizada no *Bagging*.
- Se $k = 1$, a característica de divisão é selecionada aleatoriamente.
- A recomendação típica na literatura para o valor de k é o logaritmo do número de características, isto é, $k = \log_2 d$.

O sucesso do *Random Forest* deve-se à forma como o modelo gera diversidade entre os aprendizes base (as árvores individuais). Ele combina a diversidade baseada na amostra, herdada do *Bagging*, com a diversidade baseada na seleção de características. Essa combinação amplia as diferenças entre as árvores individuais, resultando em um conjunto com maior capacidade de generalização (ZHOU, 2021).

Em termos de desempenho, o *Random Forest* pode apresentar uma taxa de erro

inicial mais elevada quando possui poucas árvores, uma vez que a restrição no conjunto de características pode prejudicar o desempenho individual de cada árvore. No entanto, à medida que mais árvores são adicionadas ao conjunto, o erro de generalização tende a diminuir, frequentemente superando o *Bagging* (ZHOU, 2021).

Além disso, o treinamento de um *Random Forest* é, em geral, mais eficiente computacionalmente do que o *Bagging* com árvores tradicionais. Isso ocorre porque as árvores do RF avaliam apenas um subconjunto reduzido de características em cada divisão, enquanto as árvores determinísticas precisam avaliar todas as características disponíveis, o que é computacionalmente mais custoso (ZHOU, 2021).

A.1.4 *Logistic Regression*

A Regressão Logística (do inglês *Logistic Regression*) é um modelo estatístico fundamental utilizado em tarefas de classificação, apesar de conter o termo "regressão" em seu nome. Ela pertence à família dos Modelos Lineares Generalizados e surge da necessidade de adaptar modelos lineares, que predizem valores reais contínuos, para cenários em que a variável resposta é discreta, como no caso de classes binárias (0 e 1) Zhou (2021). Em vez de ajustar uma linha reta aos rótulos das classes, o que seria inadequado, o modelo busca estimar uma função que separe as classes e forneça a probabilidade de pertencimento de cada amostra a uma classe específica.

O cerne do modelo reside na utilização da função logística, frequentemente chamada de sigmoide. Em uma classificação ideal, desejar-se-ia uma função degrau que alterasse instantaneamente seu valor ao cruzar a fronteira de decisão. No entanto, tal função não é matematicamente conveniente por não ser contínua nem diferenciável (ZHOU, 2021).

A Regressão Logística substitui esse degrau rígido pela função sigmoide, que apresenta um formato em "S". Essa função mapeia qualquer valor real, produzido pela combinação linear das características de entrada, para um intervalo estritamente entre 0 e 1. Trata-se de uma aproximação suave da função degrau, com propriedades matemáticas, como a diferenciabilidade em todos os pontos, que facilitam o processo de treinamento (ZHOU, 2021).

Uma característica distintiva da Regressão Logística é sua interpretabilidade probabilística. O modelo não apenas prevê o rótulo da classe, mas também estima a probabilidade a posteriori de uma amostra pertencer à classe positiva Zhou (2021). Internamente, o modelo opera por meio do *logit* ou *log-odds*, isto é, o logaritmo da razão entre a probabilidade de classe positiva e a probabilidade de classe negativa. Dessa forma, a combinação linear das entradas ($w^\top x + b$) modela diretamente essa razão logarítmica, permitindo previsões sem a imposição de hipóteses rígidas sobre a distribuição dos dados, como a normalidade (ZHOU, 2021).

Para o treinamento do modelo, isto é, para a estimação dos pesos e do viés, a Regressão Logística utiliza o Método da Máxima Verossimilhança (*Maximum Likelihood Estimation* - MLE). O objetivo consiste em ajustar os parâmetros de modo a maximizar a probabilidade do modelo prever corretamente os rótulos observados no conjunto de treinamento (ZHOU, 2021).

Do ponto de vista da otimização, esse procedimento equivale à minimização de uma função de custo, correspondente à log-verossimilhança negativa. Uma vantagem fundamental desse modelo é que sua função objetivo é convexa, o que garante que métodos de otimização numérica, como o método do gradiente ou o método de Newton, converjam para o mínimo global, isto é, para a melhor solução possível dada a base de treinamento, sem o risco de convergência para mínimos locais indesejáveis (ZHOU, 2021).

A.1.5 *Naive Bayes Classifier*

O classificador *Naïve Bayes* é um algoritmo de aprendizado de máquina supervisionado, baseado no Teorema de Bayes. Ele é classificado como um modelo generativo, pois busca modelar a distribuição de probabilidade subjacente das classes e dos dados para realizar previsões. Sua eficiência e simplicidade o tornam popular em diversas aplicações, como a classificação de texto e o diagnóstico médico (ZHOU, 2021).

O termo "*Naïve*" (ingênuo) advém da suposição central do modelo: a independência condicional dos atributos. Para tornar o cálculo da probabilidade a posteriori viável, o algoritmo assume que, dada uma classe, os atributos do objeto são independentes entre si, isto é, a presença ou o valor de uma característica não afeta a presença ou o valor de outra (ZHOU, 2021).

Embora essa suposição raramente se sustente plenamente em dados do mundo real, ela simplifica drasticamente o problema, evitando a necessidade de calcular probabilidades conjuntas complexas, que exigiriam um volume de dados exponencialmente grande e permitindo que o modelo apresente bom desempenho mesmo com conjuntos de dados limitados (ZHOU, 2021).

O processo de "treinamento" do *Naïve Bayes* é estatístico e direto, consistindo na estimativa de dois tipos de probabilidades a partir dos dados (ZHOU, 2021):

1. **Probabilidade a priori** ($P(c)$): representa a probabilidade geral de ocorrência de cada classe, calculada com base na frequência de cada classe no conjunto de treinamento.
2. **Verossimilhança** ($P(x_i|c)$): corresponde à probabilidade condicional de cada atributo dado a classe.
    - Para atributos discretos, essa probabilidade é estimada por meio da contagem de frequências.
    - Para atributos contínuos, assume-se geralmente uma distribuição (como a Gaus-

siana), cujos parâmetros, média e variância, são estimados para cada classe.

A decisão de classificação é realizada pelo cálculo da probabilidade a posteriori para cada classe, atribuindo-se à amostra a classe associada ao maior valor dessa probabilidade, de acordo com a regra de decisão de *Bayes*.

Um desafio prático do *Naïve Bayes* é o problema da "probabilidade zero". Se um determinado valor de um atributo nunca foi observado em associação a uma classe específica durante o treinamento, a estimativa correspondente seria nula. Como as probabilidades são multiplicadas, isso resultaria em uma probabilidade final igual a zero para a classe, independentemente da contribuição das demais evidências (ZHOU, 2021).

Para contornar esse problema, emprega-se uma técnica de suavização, comumente denominada correção laplaciana. Essa técnica adiciona uma pequena constante, geralmente igual a 1, às frequências observadas, garantindo que nenhuma probabilidade seja estritamente nula e preservando a capacidade do modelo de considerar informações provenientes de outros atributos (ZHOU, 2021).

O *Naïve Bayes* destaca-se ainda por sua elevada eficiência computacional. Como as probabilidades podem ser pré-calculadas ou atualizadas de forma incremental, o modelo permite (ZHOU, 2021):

- **Predição rápida**: basta consultar as tabelas de probabilidades previamente estimadas;
- **Aprendizado incremental**: o modelo pode ser atualizado com novos dados sem a necessidade de reprocessar todo o conjunto de treinamento anterior.

A.1.6 *Classification and Regression Tree*

O *CART* (*Classification and Regression Trees*) é um algoritmo fundamental no aprendizado de máquina, amplamente reconhecido por sua capacidade de gerar árvores de decisão tanto para tarefas de classificação quanto de regressão. O modelo segue uma estratégia de "dividir para conquistar", particionando recursivamente os dados em subconjuntos mais homogêneos, de modo a construir regras de decisão interpretáveis (ZHOU, 2021).

A característica definidora do CART em tarefas de classificação é o uso do Índice Gini para selecionar as divisões dos nós, diferenciando-se de algoritmos como o ID3 e o C4.5, que utilizam medidas baseadas em entropia, como o Ganho de Informação (ZHOU, 2021).

O Índice Gini mede a impureza de um conjunto de dados. Intuitivamente, ele representa a probabilidade de que dois exemplos extraídos aleatoriamente do conjunto pertençam a classes diferentes (ZHOU, 2021).

- Quanto menor o valor do Índice Gini, maior a pureza do nó, isto é, maior a concentração de exemplos de uma única classe.
- Durante o treinamento, o algoritmo avalia todas as características candidatas e seleciona a que resulta na divisão com o menor Índice Gini ponderado.

Para lidar com atributos numéricos contínuos, o CART emprega uma estratégia de bipartição. Como não é possível criar um ramo para cada valor contínuo distinto, o algoritmo busca um limiar de corte adequado (ZHOU, 2021).

Os dados são ordenados e pontos médios entre valores adjacentes são considerados como candidatos a divisor. O conjunto de dados é então particionado de forma binária: exemplos com valores inferiores ao limiar seguem por um ramo, enquanto os demais seguem por outro. O limiar selecionado é aquele que maximiza a pureza da separação resultante (ZHOU, 2021).

Árvores de decisão tendem a crescer excessivamente, capturando ruídos dos dados de treinamento e perdendo a capacidade de generalização, fenômeno conhecido como *overfitting*. O CART mitiga esse problema por meio de técnicas de poda (*pruning*) (ZHOU, 2021):

1. **Pré-poda:** Interrompe o crescimento da árvore antes que ela se torne excessivamente complexa, avaliando se uma nova divisão proporciona ganhos reais de generalização, por exemplo, por meio da validação em um conjunto de teste.
2. **Pós-poda:** Permite que a árvore cresça completamente e, posteriormente, remove subárvores que não contribuem para a generalização, substituindo-as por nós folha. Essa abordagem tende a produzir modelos mais robustos do que a pré-poda, embora implique um custo computacional maior.

A.1.7 *Multilayer Perceptron*

O *Multilayer Perceptron* (MLP), ou rede neural *feedforward*, é um modelo de aprendizado de máquina que consiste essencialmente em uma série de modelos de regressão logística (ou linear) empilhados. A arquitetura é projetada para aprender mapeamentos complexos entre entradas e saídas por meio de camadas de processamento intermediárias (MURPHY, 2012).

A camada final da rede atua como um preditor padrão, regressão logística para tarefas de classificação ou regressão linear para tarefas de regressão, operando sobre características derivadas aprendidas pelas camadas anteriores, em vez de atuar diretamente sobre os dados brutos de entrada. A estrutura de um MLP é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída (MURPHY, 2012).

- **Camadas ocultas:** O objetivo das unidades ocultas é extrair ou construir características (*feature extraction*), aprendendo combinações não lineares das entradas originais

que são mais informativas para a tarefa final do que as entradas brutas isoladas.
- **Pesos:** A rede possui matrizes de pesos que conectam as entradas às unidades ocultas e estas às unidades de saída.

Um componente crítico do MLP é a função de ativação não linear (g) aplicada às unidades ocultas. Funções comumente utilizadas incluem a sigmoide e a tangente hiperbólica (*tanh*) (MURPHY, 2012).

- A não linearidade é fundamental, pois, sem ela, o modelo seria matematicamente equivalente a uma única camada de regressão linear, independentemente da profundidade da rede.
- O MLP possui a propriedade de aproximador universal: com um número suficiente de unidades ocultas, ele é capaz de modelar qualquer função suave com o nível de precisão desejado.

O treinamento de um MLP envolve a minimização de uma função de custo não convexa, como a entropia cruzada para classificação ou o erro quadrático para regressão (MURPHY, 2012).

- **Otimização:** Utilizam-se, em geral, métodos baseados em gradiente, como o gradiente descendente estocástico, para a busca de mínimos locais da função de erro.
- *Backpropagation*: O algoritmo de retropropagação é empregado para o cálculo dos gradientes necessários, aplicando a regra da cadeia para propagar o erro da camada de saída para as camadas anteriores, possibilitando o ajuste eficiente dos pesos em toda a rede.

Devido à sua elevada capacidade de representação e ao grande número de parâmetros, MLPs são suscetíveis ao *overfitting*. Estratégias comuns para o controle da complexidade incluem (MURPHY, 2012):

- *Early stopping*: interrupção do treinamento quando o erro no conjunto de validação deixa de diminuir;
- *Weight decay*: aplicação de uma penalidade L_2 aos pesos, incentivando valores mais baixos e a obtenção de modelos mais suaves.

A.1.8 *Stacked Autoencoder*

O *Stacked Autoencoder* (SAE) é uma arquitetura de aprendizado profundo formada pela conexão sequencial de múltiplas camadas de *autoencoders*. O modelo é projetado para aprender representações hierárquicas e progressivamente mais abstratas a partir de dados de entrada de alta dimensão. Ao utilizar a saída de um codificador como entrada para o próximo, o SAE é capaz de comprimir a informação e extrair características robustas e invariantes, essenciais para tarefas complexas como reconhecimento de padrões e classificação (MURPHY, 2012).

A estrutura do SAE baseia-se no empilhamento sequencial de *autoencoders* básicos. O funcionamento de cada camada consiste na aplicação de uma transformação não linear aos dados recebidos da camada precedente, por meio de parâmetros internos, como pesos e vieses, ajustados ao longo do processo de aprendizado. Essa composição hierárquica permite que o modelo capture progressivamente a informação, abstraindo desde primitivas locais simples nas camadas inferiores até estruturas globais mais complexas nas camadas superiores (MURPHY, 2012).

O diferencial do SAE reside em seu protocolo de treinamento, dividido em duas fases principais, que visam contornar dificuldades de otimização em redes profundas, como o problema do desvanecimento do gradiente (MURPHY, 2012):

1. **Pré-treinamento não supervisionado (*layer-wise pretraining*):** Nesta etapa, cada *autoencoder* é treinado individualmente e de forma gananciosa (*greedy*), com o objetivo de minimizar o erro de reconstrução da entrada local, sem a utilização de rótulos de classe. Esse procedimento fornece uma inicialização de parâmetros mais adequada do que a inicialização aleatória, posicionando os pesos em uma região favorável do espaço de busca.
2. **Ajuste fino supervisionado (*supervised fine-tuning*):** Após o empilhamento das camadas pré-treinadas, adiciona-se uma camada de predição no topo da arquitetura, geralmente um classificador *softmax*. A rede completa é então treinada supervisionada, utilizando algoritmos baseados em gradiente para ajustar simultaneamente todos os parâmetros em direção ao objetivo final da tarefa, como a minimização da entropia cruzada em problemas de classificação.

Para garantir a capacidade de generalização e mitigar o *overfitting*, especialmente em cenários com dados limitados ou de alta dimensão, o SAE emprega diversas técnicas de regularização. Comumente, utilizam-se penalidades L_2 nos pesos, conhecidas como decaimento de peso, bem como restrições de esparsidade nas camadas ocultas. Além disso, a própria estratégia de pré-treinamento atua como um mecanismo regularizador, orientando o modelo a aprender características que capturam a estrutura intrínseca dos dados antes de se concentrar na tarefa discriminativa (MURPHY, 2012).

A.1.9 *Deep Belief Network*

A *Deep Belief Network* (DBN) é um modelo generativo profundo composto por múltiplas camadas de variáveis latentes estocásticas. Historicamente, foi uma das primeiras arquiteturas não convolucionais a permitir o treinamento eficaz de redes profundas, superando dificuldades de otimização que limitavam os modelos anteriores. Embora seu uso tenha diminuído em favor de arquiteturas mais modernas, a DBN permanece um marco relevante no desenvolvimento do aprendizado profundo (*Deep Learning*) (HINTON; OSINDERO; TEH, 2006).

A característica definidora de uma DBN é sua estrutura de conexões mista (HINTON; OSINDERO; TEH, 2006):

- **Camadas superiores (não direcionadas):** As conexões entre as duas camadas mais profundas são não direcionadas, formando uma memória associativa, especificamente uma Máquina de Boltzmann Restrita (do inglês *Restricted Boltzmann Machine* - RBM).
- **Camadas inferiores (direcionadas):** As conexões entre todas as camadas subsequentes são direcionadas, apontando de cima para baixo, em direção à camada de dados visível.

Essa configuração permite que a rede funcione como um modelo gerador, no qual as camadas superiores definem uma distribuição de probabilidade a priori sobre os dados, enquanto as camadas inferiores transformam essa representação abstrata em dados observáveis (HINTON; OSINDERO; TEH, 2006).

O treinamento de uma DBN emprega uma estratégia de aprendizado ganancioso camada por camada, que contorna a intratabilidade associada ao treinamento conjunto de modelos profundos (HINTON; OSINDERO; TEH, 2006):

1. **Pré-treinamento não supervisionado:** A rede é construída por meio do empilhamento de RBMs. A primeira RBM é treinada diretamente sobre os dados de entrada, e as ativações de sua camada oculta são então utilizadas como dados de entrada para o treinamento da RBM subsequente. Esse processo é repetido sucessivamente, permitindo que cada camada modele as correlações da camada anterior.
2. **Ajuste fino (*fine-tuning*):** Após o pré-treinamento, os pesos da rede generativa são utilizados para inicializar uma rede neural discriminativa, como um MLP. A rede completa é então refinada para uma tarefa específica, como classificação, por meio de algoritmos de aprendizado supervisionado convencionais.

Como modelo generativo, a DBN é capaz de gerar novas amostras de dados. Esse processo envolve duas etapas principais: inicialmente, realiza-se a amostragem de Gibbs nas camadas superiores não direcionadas até que o sistema atinja o equilíbrio. Em seguida, executa-se uma passagem única e determinística, conhecida como amostragem ancestral, ao longo das conexões direcionadas descendentes, para gerar os valores correspondentes na camada visível (HINTON; OSINDERO; TEH, 2006).

A.1.10 *Recurrent Neural Network*

As *Recurrent Neural Networks* (RNNs) constituem uma família de redes neurais especializadas no processamento de dados sequenciais. Elas são projetadas para processar sequências de comprimento variável, distinguindo-se das redes *feedforward* tradicionais pela capacidade de manter uma memória de eventos passados por meio de ciclos em seu grafo computacional (GOODFELLOW; BENGIO; COURVILLE, 2016).

O conceito central das RNNs é o compartilhamento de parâmetros ao longo do tempo. Ao utilizar os mesmos pesos em cada passo temporal, a rede consegue generalizar o aprendizado para sequências de diferentes tamanhos e posições temporais (GOODFELLOW; BENGIO; COURVILLE, 2016).

A operação da rede pode ser compreendida por meio do processo de "desdobramento" (*unfolding*) do grafo recursivo. Esse procedimento transforma a estrutura cíclica em uma cadeia repetitiva de operações, na qual o estado oculto atual depende do estado oculto anterior e da entrada corrente. Do ponto de vista matemático, o estado oculto atua como um resumo compactado de toda a sequência de entradas processada até o instante considerado (GOODFELLOW; BENGIO; COURVILLE, 2016).

Existem diferentes padrões de projeto para RNNs, que variam conforme a natureza das conexões recorrentes (GOODFELLOW; BENGIO; COURVILLE, 2016):

- **Recorrência oculta-para-oculta:** Quando as conexões ocorrem entre unidades ocultas, a rede atinge seu maior poder de representação, sendo teoricamente capaz de simular qualquer algoritmo computacional, conforme o modelo da Máquina de Turing.
- **Recorrência saída-para-oculta:** Quando a recorrência ocorre apenas a partir da saída anterior, a rede apresenta menor capacidade de modelar dependências de longo prazo, porém ganha eficiência de treinamento por meio da paralelização.

O treinamento de RNNs utiliza o algoritmo *Backpropagation Through Time* (BPTT), que propaga os erros ao longo da sequência temporal desdobrada. Devido à grande profundidade temporal envolvida, esse processo tende a ser computacionalmente custoso (GOODFELLOW; BENGIO; COURVILLE, 2016).

Para arquiteturas que utilizam realimentação da saída, frequentemente aplica-se a técnica de *Teacher Forcing*. Durante o treinamento, a rede é alimentada com os valores reais do passo temporal anterior, em vez de suas próprias predições, o que contribui para a estabilidade e aceleração do processo de aprendizado (GOODFELLOW; BENGIO; COURVILLE, 2016).

A.1.11 *Long Short-Term Memory*

O *Long Short-Term Memory* (LSTM) é uma variação avançada das Redes Neurais Recorrentes (do inglês *Recurrent Neural Networks*, RNNs), projetada especificamente para superar as limitações das RNNs tradicionais no aprendizado de dependências de longo prazo. Desde sua introdução, o modelo consolidou-se como uma ferramenta fundamental em tarefas de processamento sequencial complexo, como tradução automática e reconhecimento de fala (GOODFELLOW; BENGIO; COURVILLE, 2016).

A arquitetura distingue-se por sua capacidade de preservar o fluxo de gradientes ao

longo de extensos intervalos temporais, facilitando o treinamento em sequências nas quais a informação relevante está distante do ponto de predição atual Goodfellow, Bengio e Courville (2016). A principal inovação do LSTM consiste na substituição de unidades recorrentes simples por células de memória dotadas de uma estrutura interna mais elaborada. O núcleo dessa estrutura é o emprego de portas (*gates*), responsáveis por controlar o fluxo de informação no interior da célula (GOODFELLOW; BENGIO; COURVILLE, 2016).

Diferentemente das redes recorrentes tradicionais, que utilizam pesos fixos em seus laços de recorrência, o LSTM emprega conexões autorrecorrentes (*self-loops*) cujos pesos são ajustados dinamicamente em função do contexto da rede. Essa característica permite que a escala temporal de integração da informação varie conforme a sequência de entrada, permitindo que o modelo decida quando reter ou descartar informações (GOODFELLOW; BENGIO; COURVILLE, 2016).

O funcionamento da célula LSTM é controlado por três mecanismos principais, denominados portas, que utilizam funções de ativação sigmoide para regular a passagem de dados, com valores variando entre 0 e 1 (GOODFELLOW; BENGIO; COURVILLE, 2016):

1. **Porta de esquecimento (*forget gate*):** Responsável por determinar quais informações do estado anterior da célula devem ser descartadas, permitindo que a memória seja atualizada quando o contexto se altera ou quando informações passadas se tornam irrelevantes.
2. **Porta de entrada (*input gate*):** Regula a quantidade de nova informação proveniente da entrada atual que será incorporada ao estado da célula, atuando em conjunto com a porta de esquecimento na atualização do estado interno.
3. **Porta de saída (*output gate*):** Controla qual parcela do estado interno da célula será transmitida para a saída da rede e para o próximo passo temporal, permitindo que informações relevantes sejam mantidas internamente sem necessariamente serem expostas em todos os instantes.

A.1.12 *Gated Recurrent Unit*

O *Gated Recurrent Unit* (GRU) é uma arquitetura de rede neural recorrente pertencente à família das RNNs com portas (*Gated RNNs*). Desenvolvido como uma alternativa simplificada ao LSTM, o GRU tem como objetivo mitigar o problema do desaparecimento do gradiente e capturar dependências de longo prazo em sequências, permitindo que a rede ajuste dinamicamente o fluxo de informação ao longo do tempo (GOODFELLOW; BENGIO; COURVILLE, 2016).

A principal inovação do GRU em relação ao LSTM reside na redução da complexidade de seus mecanismos internos. Diferentemente do LSTM, que possui portas dedicadas para entrada, saída e esquecimento, o GRU unifica parte dessas funcionalidades em um

número menor de componentes (GOODFELLOW; BENGIO; COURVILLE, 2016).

Uma característica definidora do GRU é o fato de uma única unidade de porta ser responsável por controlar simultaneamente o fator de esquecimento e a decisão de atualização do estado. Essa simplificação resulta em um modelo com menor número de parâmetros, mantendo, entretanto, capacidade semelhante de modelagem temporal (GOODFELLOW; BENGIO; COURVILLE, 2016).

O fluxo de informação dentro de uma unidade GRU é controlado por duas portas principais (GOODFELLOW; BENGIO; COURVILLE, 2016):

1. **Porta de atualização (*update gate*):** Atua como um mecanismo de integração condicional, determinando, para cada dimensão do estado, se a rede deve preservar o valor do estado anterior ou substituí-lo por uma nova informação. Esse mecanismo permite que informações relevantes sejam mantidas por longos períodos.
2. **Porta de redefinição (*reset gate*):** Controla a quantidade de informação do estado passado utilizada no cálculo do estado candidato. Ao permitir que a rede ignore partes do passado irrelevantes para o contexto atual, essa porta introduz uma não linearidade adicional essencial para modelar transições de estado mais complexas.

Estudos empíricos que comparam diferentes variantes de arquiteturas recorrentes indicam que tanto o GRU quanto o LSTM apresentam desempenho elevado. Em geral, não há uma superioridade clara de um modelo sobre o outro em todas as tarefas, sendo comum que ambos superem outras variantes arquiteturais e se consolidem como referências na modelagem de sequências (GOODFELLOW; BENGIO; COURVILLE, 2016).

APÊNDICE B – Tabela resumo dos trabalhos de revisão bibliográfica

A seguir, descrevem-se os conteúdos de cada coluna da tabela:

- **Proposta**: técnicas e metodologias empregadas em cada estudo (modelos de *Machine Learning*, métodos estatísticos, abordagens híbridas, entre outros);
- **Entrada (x_t)**: variáveis de entrada utilizadas nos experimentos, incluindo, quando aplicável, os Indicadores Técnicos (IT) derivados dessas variáveis;
- **Saída ($\hat{y}_t$)**: variáveis de saída previstas pelos modelos, tais como preços, retornos, volatilidade ou sinais de negociação;
- **Seleção**: métodos de seleção ou redução de características adotados pelos autores, quando aplicável;
- **Horizonte**: horizontes de previsão considerados (minuto, fim do dia — EOD, semanal, mensal, entre outros);
- **Ações ou índices**: ativos financeiros ou índices de mercado utilizados nos experimentos;
- **Intervalos**: períodos temporais dos dados analisados e, quando disponíveis, a quantidade de amostras e a proporção dos dados utilizados para treinamento (T) e validação (V);
- **Métricas**: métricas de desempenho utilizadas na avaliação dos modelos propostos.

Tabela 34 – Resumo dos trabalhos relacionados.

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
			Agarwal et al. (2025)					
SVMD-LSTM	Close	-	Close	-	1 dia	HSI S&P 500 SENSEX WTI	13/12/2007 – 12/12/2017 13/12/2007 – 12/12/2017 01/01/2013 – 30/12/2022 02/01/2008 – 26/12/2013 2465 amostras 2518 amostras 2476 amostras 1510 amostras 90% T / 10% V	*RMSE* *MAE* *MAPE* *R²* *CID*
			Safari e Ghaemi (2025)					
Neuro-Fuzzy + BiLSTM	OHLC, Adj Close, Volume	-	OHLC	-	1 dia	JPM AMZN TSLA	11/10/1985 a 04/07/2023 16/05/1985 a 04/07/2023 02/07/2010 a 04/07/2023 11.024 amostras 6.684 amostras 3.383 amostras 80%T / 20%V	*RMSE* *MAE* *MAPE* *RMSPE* *MSE* *R²* *F1* *FPS*
			Ge (2025)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
MEMD-AO-LSTM	OHLC, Volume	-	Close	-	1 dia	S&P 500 CSI300 Nasdaq100 Nikkei225 FTSE DAX SSE KOSPI	01/01/2013 a 01/01/2022 aprox. 2500 amostras por ativo 80%T / 20%V	*MAPE* *RMSE* *MAE* R^2
Uçkan (2024)								
LSTM *CNN* *BiLSTM* *GRU*	. Date, OHLC, Adj Close, Volume	WMA, EMA, RSI, CMO, Williams %R, ROC, HMA, TEMA, ADX, PLine	Close	PCA	1 dia	ASELS, TUPRS, THYAO, SISE, FROTO	01/01/2014 a 01/01/2024 2317 amostras 80%T / 20%V	*MSE* *MAE* *MAPE* R^2
Tashakkori et al. (2024)								
RNN LSTM	Close	-	Close	-	1 dia	CAC40	01/01/2010 a 01/01/2021 1394 amostras 80%T / 20%V	*MSE* *RMSE*
Zheng et al. (2024)								
RF	OHLC	RSI, Stochastic Oscillator, Williams %R, MACD, Price Rate of Change, OBV	categorias: alta, lateralização e baixa.	-	30/60/90 dias	Apple General Eletric Samsung	06/07/2013 a 01/01/2024 7000 amostras Treinamento via Out-of-Bag (OOB)	OOB Accuracy *Tempo de processamento*
Al-Khasawneh et al. (2024)								

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t) Valores Brutos	Indicadores Técnicos	Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
LSTM *BiLSTM* *ARIMA* *SARIMA* *ARMA* *ARCH/GARCH*	Close, log-returns, High-Low Spread (%), Open-Close Spread e Volatilidade atual	-	Weekly Vol PriceDir CloseTrend	-	1 dia	PSX100	01/01/2009 a 07/07/2021 2980 amostras 75%T / 25%V	*RMSE* *MSE* *RMSPE*
			Shaban, Ashraf e Slama (2024)					
BiGRU-LSTM	OHLC, Volume	SMA	Close	-	10 minutos 30 minutos	IBM	25/03/2020 a 06/04/2022 216.833 amostras 80%T / 20%V	*MAPE* *MSE* *MAE* R^2
			Abolmakarem et al. (2024)					
MLP LSTM	Lagged log-returns (OHLC)	RSI, Momentum, TR, ATR, Parabolic SAR, CCI, SMA, WMA, HMA, ADX, William's %R, %K, %D	Trend, PK, TRU, DBP, DBT, DBPT, PKHe, TRUDe Delta, Ret. (daily return, %)	-	1 dia	FTSE100: BP BHP GSK VOD	01/01/2000 a 31/12/2020 5283 amostras 70%T/15% V/15%Te	Accuracy Precision F1 Recall RMSE MSE R^2
			Safari e Badamchizadeh (2024)					
SoLTD-BiLSTM	OHLC, Adj Close, Traded Volume	-	OHLC, Adj Close e Volume	-	1 dia	AMZN	15/05/1997 a * 19.971 amostras 80% T / 20% V	R^2 *MAPE* *RMSPE*
			Mutinda e Langat (2024)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
GARCH-LSTM GARCH-GRU GARCH-Transformer	OHLC, Adj Close, Volume, Volatilidade condicional	-	Close	-	1 dia	Airtel	28/06/2019 a 08/05/2020 amostras não informadas 67% T / 33% V	R^2 *MAPE* *RMSE* *MAE*
			Pan, Tang e Wang (2024)					
MULTI-GARCH-LSTM	OHLC, volume de negociação, open interest, turnover, variação de preço	-	Close	-	1 dia	SHFE	01/01/2013 a 30/09/2023 2608 amostras 80% T / 20% V	R^2 *MAPE* *RMSE* *MAE* *MSE*
			Botunac, Bosna e Matetić (2024)					
LSTM	OHLC	*TEMA* *MOM* *MACD* *P-MA* *SMA* *WMA*	Close	-	1 dia	*S&P 500* *DIA* *AAPL* *MSFT* *TSLA* *NVDA*	01/01/2015 a 01/10/2023 amostras não informadas 70%T / 30%V	Retorno Acumulado(%) *MAE* *MSE*
			Bhanja e Das (2024)					
BSEI k-NN SVR CNN-autoencode MNB GBM ADA DML	OHLC	*MA* *SMA* *RSV* *%K* *%D* *ROC* *RSI*	Future Trend (1,0,-1)	-	1 dia 10 dias	BSE-SENSEX Nifty50	01/01/1991 a 31/12/2019 01/01/1996 a 31/12/2019 amostras não informadas 70% T / 30% V	*Accuracy* *Precision* *Recall* *F1*

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
			Das e Others (2024)					
LSTM GRU	OHLC	Bollinger Volatilidade Rolling	Close	-	1 dia	*AAPL* *JNJ* *JPM* *CVX* *BTC* *S&P 500*	1998 a 2023 amostras não informadas	*R²* *RMSE* *MSE* *MAE* *MAPE*
			Yasin, Paschke e Qundus (2024)					
DQN	OHLC, Volume, DAX e Google Trends	-	Direção do preço (binary up/down)	-	1 hora 1 dia 1 sem	ADS BASF ALV	01/01/2017 a 31/12/2021 amostras não informadas 60%T / 20% V 20%Test	*MDA* *RMSE*
			Yenireddy e Others (2024)					
LSTM LR SVM	OHLC e Volume	Momentum Volatilidade SMA	Close	-	1 dia	S&P 500	01/01/2010 a 31/12/2022 3391 amostras 80%T / 20%V	*RMSE* *R²* *MAE*
			Phuoc e Others (2024)					
LSTM	OHLC e Volume	SMA MACD RSI	Close	-	1 dia	VN-Index VN30	XX/XX/XX a 01/04/2021 (varia por ticker) amostras não informadas 90%T 10%V	*Aj (accuracy por ticker)*
			Gupta e Jaiswal (2024)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
ARIMA LSTM RNN Full CNN	OHLC, Volume	-	Close	-	1 dia	S&P 500	01/12/2015 a 31/01/2019 amostras não informadas	*MAE*
			Saxena, Sharma e Gupta (2023)					
ARIMA LSTM	Close	-	Close	-	1 dia	HLC IHG MII	01/01/1999 a 31/12/2021 amostras não informadas 80%T / 20%V	RMSE
			Vullam et al. (2023)					
MMGAN-HPA	OHLC	-	Close Daily Return	XGBoost PCA	1 dia	TCS BHEL TATASTEEL AXISBANK	Nenhuma informação	Correlation *MAE* *MSE*
			Zaichenko et al. (2023)					
TFT RoBERTa	OHLC, Adj close, Volume e Sentiment Score	ATR EWMA	Close	-	3 dias 5 dias 15 dias 32 dias	AMZN AAPL GOOGL TSLA MSFT	01/01/2015 a 31/12/2020 Dataset Twitter: 3.717.964 tweets 80%T / 20%V	*RMSE* *MAPE* *MAE* R^2 *SMAPE*
			Yang, Wang e Li (2022)					
LASSO-LSTM *LSTMTI* *RF* *SVM* *LSTM*	OHLC, Volume, mais quatro indicadores de sentimento	ADX, ATR, BBands, CCI, CMF, MACD, RSI, SAR, SMA/EMA/DEMA/WMA, OBV, MFI, ROC, DPO, KST, HMA, SMI, Chaikin AD, Donchian, GMMA	Direção do movimento do preço (binário)	LASSO	1 dia	AAPL MNST BAC	01/01/2012 a 31/12/2020 aprox. 1000 amostras por ativo 75%T / 25%V	*Accuracy* *Kappa* *AUC* *Sensitivity* *Specificity* *Confusion matrix*
			Padhi et al. (2022)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
Ensemble incremental (PA+Perceptron)	OHLC e Volume	*SMA, Parabolic SAR, Aroon, Balance of Power (BOP), Directional Movement Index (DMI), Chaikin A/D, OBV, True Range e COS*	Direção do preço (binário)	*mean-variance VIF SMOTE*	1 dia	*DHR NVO UNH DPHL AFX MRK CVS EWL*	séries históricas até 24/11/2021 amostras não informadas 75% T / 25% V	Precision AUC *Accuracy* Recall F1 Hamming Loss
Chhajer, Shah e Kshirsagar (2022)								
MLP LSTM CNN DNN SVM SVR LSTM-CNN	OHLCV	*SMA, RSI, MACD, Bollinger Bands, ATR, OBV, Chaikin A/D, Parabolic SAR, CCI, DMI, ROC*	Close Direção do preço (binário)	-	1 dia	S&P 500 DJIA NASDAQ CSI300 Nikkei	01/01/1991 a 31/12/2020 amostras não informadas 80%T / 20%V	Accuracy Precision Recall AUC Kappa Confusion Matrix *MSE RMSE MAE MAPE* R^2
Huang, Capretz e Ho (2022)								
FNN RF ANFIS	21 variáveis fundamentais	-	Relative quarterly return	RF	1 trimestre	S&P100	Q1-1996 a Q4-2017 88 observações 60%T / 20%V 20%Test	*Portfolio Score MQRR CRR RMSE*
Singh et al. (2022)								

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
LR *MLP* *SVM* *SGD* *kNN* *DT* *ADA* *RF*	OHLC	-	Close	-	1 dia	NIFTY50	22/04/1996 a 16/04/2021 6220 amostras 80%T 20%V	*MSE* *RMSE* *MAE* *R²* *CVRMSE*
			Chandra e He (2021)					
Bayes-FNN *FNN-Adam* *FNN-SGD*	Close Price Adj	-	Close	-	1-5 dias	MMM 600118.SS CBA.AX DAI.DE	01/01/2012 a 01/07/2020 aprox. 400 observações por ativo 80%T / 20%V	*RMSE*
			Mehta, Pandya e Kotecha (2021)					
LSTM NB SVM/SVC MLP DT	News / social media: (date, author, title, content) e OHLC	SMA	Classificação binária, Close e Open	-	1 dia 10 dias 15 dias 30 dias 90 dias 180 dias	BSE	01/01/2009 a 31/12/2019 aprox. 82000 amostras	*Accuracy* *Precision* *Recall* *F-score*
			Singh e Khushi (2021)					
RF MLP kNN SVM DT NB	28 features (combinação de fundamentais, técnicos, recomendações de analistas e derivados)	*Market Cap.* *Historical Market Cap.* *Short Interest Ratio.*	Classificação ternária: Buy (2), Hold (1), Sell (0)	*PCA*	1-10 dias	S&P 500	04/04/1999 a 04/04/2019 amostras não informadas 70% T / 30% V	Recall F1 Precision
			Farahani e Hajiagha (2021)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
ANN / MLP+GA BA SSO ARMA ARIMA	OHLC, Volume	ADL, CMF, MFI, TP, RSI, ROC, Bollinger Bands (Upper/Lower), MP, ATR, CCI, DX, SMA e EMA	Close	*GA*	1 dia	S&P 500 DAX FTSE100 NASDAQ DJI	04/07/2018 a 04/07/2020 aprox. 438 amostras por ativo 70% T / 30% V	MSE MAE correlação AIC R^2
Nabipour et al. (2020)								
LSTM DT BAG RF ADA GB XGB RNN	OHLC	SMA, WMA, Momentum, Stochastic K%, Stochastic D%, RSI, MACD (Signal), Larry Williams R%, Accumulation/Distribution (A/D), CCI)	stock-group index value	-	1 dia 1 dia 2 dias 5 dias 10 dias 15 dias 30 dias	TSE.IR	01/11/2009 a 01/11/2019 2600 amostras 70% T / 30% V	MSE MAE RRMSE MAPE
Pang et al. (2020)								
ELSTM AELSTM DBN DBN-MLP MLP	OHLC, Volume	Daily amplitude Volume Fluctuation	Close Direção(up/down)	-	1 dia	SSE Index Sinopec TMSE TBEA	01/01/2006 a 19/10/2016 amostras não informadas 70% T / 20% V 10% Test	MSE Data Accuracy
Zhang et al. (2019)								
GAN LSTM MLP	OHLC, Volume e Turnover Rate	SMA	Close	-	1 dia	S&P 500 SSE IBM MSFT PAICC	01/01/1998 a 01/01/2018 aprox. 5000 amostras por ativo 90% T / 10% V	MSE MAPE RMSE Average Return
Hiransha et al. (2018)								

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
MLP RNN LSTM CNN ARIMA	Close	-	Close	-	10 dias 20 dias 30 dias 40 dias 400 dias	TATAMOTORS BAC	01/01/1996 a 30/06/2015 4861 amostras 05/10/2007 a 30/06/2017 2430 amostras T e V não informadas	MAPE
			Ince e Trafalis (2017)					
TWSVM MPM KFDA	OHLC, Volume	EMA(10), EMA(50), Stochastic %K, Price Rate of Change, RSI, Accumulation/Distribution Oscillator, MACD, Williams %R, High Price Acceleration, Disparity5, Disparity10	direção do movimento diário (binário)	*ICA*	1 dia	S&P 500 NASDAQ100 DJIA	03/08/2007 a 23/05/2015 2037 amostras 80% T / 20% V	Hit Ratio Excess Return Sharpe Ratio Accuracy
			Qiu, Song e Akagi (2016)					
ANN / BPNN GA+BPNN SA+BPNN	71 variáveis iniciais	-	Retorno Mensal	*Fuzzy Surfaces*	1 mês	Nikkei225	01/11/1993 a 01/07/2013 237 amostras 70% T / 30% V	MSE
			Attigeri et al. (2015)					
Sentiment Analysis + LR	OHLC, Volume, Adj. Close, indicadores sintéticos de sentimento.	-	Tendência Close (binário)	*Correlação*	1 dia	Company1 Company2	Intervalo não mencionado 3900 amostras por ativo 50% T / 50% V	Accuracy Confusion Matrix
			Barak e Modarres (2015)					
DT DTNB MLP SVM kNN	44 variáveis	-	Real return N-S Risk	*HFS*	1 dia	TSE.IR	01/01/2003 a 31/12/2012 1963 amostras 90% T / 10% V	Accuracy Sensitivity Specificity
			Ievsieieva et al. (2024)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
VAR GARCH ARIMA	Stock Returns, Market Index e Volatility Measure	-	Retorno Mensal	-	1 mês	Empresa média índice setorial MI	01/01/2019 a 31/12/2023 60 amostras	AIC BIC Ljung-Box RMSE MAE
			Daryl et al. (2021)					
SARIMA SARIMAX	Close	-	Close Mensal	-	1 mês	AAPL	01/01/2010 a 31/12/2019 2518 amostras	RMSE MAPE BIC AIC HQIC
			Idrees, Alam e Agarwal (2019)					
AR MA ARMA ARIMA	OHLC, Volume	-	Close Mensal	-	1 mês	NIFTY SENSEX	01/01/2012 a 31/12/2016 amostras não informadas 80% T / 20% V	ADF log-likelihood BIC AIC AICc
			Adebiyi, Adewumi e Ayo (2014)					
ARIMA	OHLC	-	Close	*GA*	1 dia	NOKIA ZenithBank	25/04/1995 a 25/02/2011 3990 amostras 03/01/2006 a 25/02/2011 1296 amostras	BIC ADF R^2
			Dutta, Bandopadhyay e Sengupta (2012)					
LR (binária)	NS, CEPS, BV, PECEPS, PE, PBIDTS, PEBV e SNA	-	Retorno Anual da Ação	-	1 ano	NIFTY	01/01/2005 a 31/12/2008 aprox. 4 amostras por ativo	Accuracy Matrix Confusion Wald statistics Omnibus test Hosmer–Lemeshow
			Merh, Saxena e Pardasani (2010)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
ANN+ARIMA	OHLC	-	Close	-	1 dia	BSE30 BSEIT BSE100 CNX Nifty S&P	16/04/2004 a 16/04/2009 aprox. 1218 amostras	RMSE MAPE AAE MSPE ACF/PACF
			Faria et al. (2009)					
ANN AES	Close	-	Close	-	1 dia	IBOV	01/09/1998 a 30/04/2007 2368 amostras 90% T / 10% V	RMSE NTEND
			Engle, Focardi e Fabozzi (2008)					
ARCH/GARCH	Retornos	-	Volatilidade condicional VaR	-	1 dia	S&P 500 NASDAQ DJIA	23/03/1990 a 23/03/2000 aprox. 438 amostras por ativo	AIC BIC log-likelihood Ljung-Box VaR
			Lee, Yoo e Jin (2007)					
BPNN SARIMA	Close e Retorno	-	Close e Retorno	-	1 semana	KOSPI	04/01/1999 a 29/05/2006 390 amostras 80% T / 20% V	FEt MAE MSE
			Zhang (2024)					
LR+PT	OHLC e prospect index	-	Close Direção (binária)	-	1 dia	S&P 500 BTC	1950 a 2015 16.590 amostras (S&P 500) 129.126 amostras (BTC)	RMSE MAE Accuracy R^2
			Puppo et al. (2022)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
TODIM	OHLC, Volume	SMA, Bollinger, MACD, RSI, Momentum, Stochastic	recomendação de compra/venda	-	1 dia	BBDC4 ITUB4 BRKM5 CMIG4 GGBR4 VIVT4 PETR4 VALE3	03/01/2002 a 31/12/2017 3961 amostras por ativo	Profit Sharpe Ratio Max. DD Calmar Expect
			Wang et al. (2022)					
3WD+CPT+Fuzzy	Retornos futuros	-	proporção de investimento ótima e final	-	60 dias	NYSE DJIA	01/01/2019 a 31/03/2019	VaR Highest return Lowest return Medium return
			Lv et al. (2023)					
TODIM+(12 algoritmos de ML)	OHLC, Volume	44 TI	Trading Signal e Modelo de Trading Ótimo	-	1 dia	S&P 500 CSI300	01/01/2017 a 30/06/2020 aprox. 882 amostras por ativo Walk-Forward Analysis	WR ARR ASR MDD CAR SOR MSR ART
			Jain, Sahay e Nupur (2025)					

Tabela 34 – Resumo dos trabalhos relacionados (continuação)

Proposta	Entrada (x_t)		Saída ($\hat{y}_t$)	Seleção	Horizonte	Ações ou índices	Intervalos	Métricas
	Valores Brutos	Indicadores Técnicos						
REM Clustering TODIM BN- XGBoost TLBO	OHLC+12 indicadores financeiros	-	Retornos diários futuros e distribuição de investimento ótima	-	1 dia	Nifty100	01/01/2018 a 25/07/2023 aprox. 1386 amostras	MSE MAE RMSE R^2 Retorno do Portfólio Risco do Portfólio CVaR

APÊNDICE C – Tabelas de Desempenho por Métrica e Cenário

Tabela 35 – Taxa de Retorno Anualizado (ARR) no cenário InS-TraD

Modelo	BM	CC	CNC	COM	FIN	GCS	OGB	UTI
Index	0,2161	0,2257	0,2219	0,2502	0,2403	0,2405	0,2200	0,2339
BAH	**0,3137**	0,3686	0,2153	1,3505	**0,5840**	**0,6534**	0,3199	**0,5079**
LR	0,1766	0,3082	0,2107	0,1027	0,5314	0,4463	0,2757	0,3752
SVM	0,1675	**0,3831**	**0,2635**	1,3692	0,5212	0,5440	**0,5382**	0,3791
NB	0,2134	0,2596	0,2072	0,6490	**0,5813**	0,2347	0,3452	0,3169
CART	0,1824	0,3317	0,2207	**1,4413**	0,4813	0,5534	0,3580	0,3502
RF	0,2441	0,3121	0,2517	0,7610	0,5752	0,4672	0,3671	0,3882
XGB	0,2057	0,3062	0,1867	0,4021	0,5692	0,5347	0,0610	0,3490
MLP	0,1820	0,2956	0,1313	1,3789	0,3124	**0,5852**	0,1148	0,3344
DBN	0,1700	0,3014	0,2437	0,0710	0,4945	0,5347	0,0525	0,3668
SAE	0,1926	0,2651	0,2279	1,3817	0,5322	0,5683	0,0154	0,3954
RNN	0,1877	0,2556	0,2500	1,3954	0,4718	0,4914	0,3040	**0,4077**
LSTM	**0,2448**	0,3009	0,2020	1,4401	0,4803	0,5038	0,1905	0,3780
GRU	0,2149	0,2716	0,1362	1,3312	0,5037	0,4602	0,1716	0,3904

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 36 – Taxa de Retorno Anualizado (ARR) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,3679	−0,4221	−0,2825	−0,2835	−0,3054	−0,2684	−0,3058	−0,2708
BAH	−0,0051	0,0476	0,3110	0,8172	−0,2758	0,2116	0,2475	0,0328
LR	0,0884	0,0788	0,2283	0,2432	−0,2979	0,4083	−0,2073	0,0516
SVM	0,1090	0,0384	0,2016	**1,2868**	−0,1990	0,4027	0,3283	**0,2610**
NB	0,0610	0,0379	−0,0991	1,0106	−0,1443	**0,4744**	0,1620	−0,0406
CART	−0,0249	**0,1389**	0,3707	0,1394	−0,1023	0,0475	−0,0175	−0,0384
RF	−0,1027	−0,0469	0,0372	0,8017	−0,1707	0,2580	0,3384	0,0929
XGB	0,0385	0,0770	0,0093	0,2743	−0,0866	0,1031	0,3117	−0,0061
MLP	−0,1127	0,1210	0,3916	0,8134	−0,1642	0,1368	−0,0212	−0,0080
DBN	−0,1306	0,0366	**0,3983**	0,3786	−0,1471	0,2232	0,1667	0,0086
SAE	−0,1153	0,0679	0,3096	0,6893	−0,1911	0,1755	−0,0182	0,0519
RNN	0,1006	0,0515	0,2101	0,6238	−0,0681	0,2044	0,1580	0,1118
LSTM	0,0279	−0,0528	0,3152	0,1189	−0,2660	0,1310	−0,0534	0,0449
GRU	**0,1869**	0,0480	0,0958	0,4508	**−0,0526**	0,3225	**0,4620**	0,1120

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 37 – Taxa de Retorno Anualizado (ARR) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	0,1984	0,2203	0,2078	0,2224	0,2262	0,2212	0,2571	0,2336
BAH	**0,3043**	0,2800	0,2830	0,2058	0,3522	**0,3928**	0,1847	**0,2548**
LR	0,2279	0,2598	0,1658	0,2539	0,3593	0,2227	−0,1075	0,1456
SVM	0,1827	0,1886	0,0903	0,2590	0,3849	0,1554	0,2229	0,1124
NB	0,0755	0,2047	0,2175	0,3069	0,3078	0,1855	0,1492	0,0931
CART	0,2021	0,1335	0,1890	0,2820	0,2098	0,2766	**0,2826**	0,1913
RF	0,1946	0,1438	0,1644	**0,4098**	**0,4036**	0,2321	0,2228	0,1763
XGB	0,2494	0,1571	0,2888	0,2946	0,3005	0,1278	0,2225	0,1716
MLP	0,2150	**0,2819**	0,0900	0,0947	0,3226	0,2524	0,2107	0,0306
DBN	0,1735	0,0863	0,1219	0,0000	0,3097	**0,3178**	0,1317	0,0900
SAE	0,2558	0,2019	−0,0142	0,0803	0,3137	0,2324	0,1308	0,1719
RNN	0,2121	0,1859	**0,3333**	0,1311	0,2890	0,1522	−0,0206	**0,2433**
LSTM	**0,2963**	0,1072	0,2064	0,0759	0,2311	0,1370	0,1804	0,1977
GRU	0,2249	0,0733	0,1556	0,2544	0,2402	0,1292	0,1680	0,1299

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 38 – Taxa de Retorno Anualizado (ARR) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,2875	−0,3058	−0,3058	−0,3058	−0,3058	−0,3058	−0,3058	−0,3058
BAH	−0,0971	−0,5206	−0,0539	0,2294	3,3816	0,1334	−0,3818	−0,0042
LR	−0,2106	−0,3805	−0,5122	0,0941	−0,2298	−0,0446	−0,0408	−0,0930
SVM	0,1582	−0,5435	−0,0511	0,6081	−0,5942	0,1044	0,1930	0,2209
NB	**0,6100**	−0,4254	0,0886	0,1755	0,0201	−0,0100	−0,5001	0,0176
CART	−0,0741	−0,6368	−0,2632	0,4302	−0,2384	−0,0207	0,1750	0,1749
RF	0,0654	−0,8066	0,0106	**1,3388**	−0,7025	−0,1048	−0,0404	**0,2619**
XGB	−0,1055	−**0,0754**	−0,2307	0,6405	3,5864	−0,2276	**0,5490**	0,0314
MLP	−0,4444	−0,4734	0,0704	0,2294	3,4940	**0,4071**	−0,5542	−0,0191
DBN	−0,2580	−0,0843	−0,0393	0,0000	3,4610	0,3712	−0,5130	−0,0829
SAE	−0,0376	−0,1741	**0,2471**	0,3495	3,5308	0,2644	−0,2291	−0,1062
RNN	0,2005	−0,2343	−0,0381	0,9642	**3,7819**	0,2117	0,3163	0,0386
LSTM	−0,1482	−0,1531	−0,1789	0,2423	3,5286	−0,1393	−0,4241	−0,2693
GRU	0,0464	−0,1154	0,1233	1,0036	3,6341	0,0828	−0,4515	0,2026

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 39 – Tempo Médio de Recuperação (ART) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**15,3230**	**16,6136**	**15,5285**	**13,6824**	**15,3539**	**15,2903**	**15,9100**	**15,1376**
BAH	42,9301	69,1622	73,4370	63,5247	38,7546	87,2275	65,0744	19,6403
LR	52,5386	74,8499	101,9115	63,6921	54,0544	66,6231	56,3487	28,1934
SVM	76,6931	69,7371	98,2999	77,0133	40,8744	**53,3544**	**42,5909**	25,1129*
NB	54,8466	**63,5822**	69,9767	60,5897	43,4950	74,9386	59,0032	26,0503
CART	53,7705	95,0880	119,8576	57,5467	43,8850	71,9504	90,7218	31,5241
RF	50,2386*	67,6557	70,1286	63,8176	39,9676	61,7261	87,7546	27,6836
XGB	50,8564	85,2854	84,5635	57,7615	53,9995	72,7369	120,9011	31,2471
MLP	57,0163	80,9774	52,9694	88,4038	39,2828	87,0945	136,5807	47,7317
DBN	66,2087	70,5680	**52,8579**	**41,1100**	41,4533	95,2855	93,7643	30,6816
SAE	55,0290	87,2620	59,5621	90,3857	**37,8835**	105,5205	137,0795	29,6638
RNN	72,2587	83,7843	91,4025	48,4485	59,7474	111,4395	78,0319	41,4324
LSTM	50,2896	79,8230	135,6834	48,5719	53,7312	102,1968	133,1500	32,8697
GRU	67,4750	98,9742	143,2131	98,7264	59,0773	112,4134	112,1014	28,4217

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 40 – Tempo Médio de Recuperação (ART) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	37,5600	38,8850	38,4000	38,4000	39,9557	39,7944	38,6667	37,7688
BAH	33,0041	39,3638	31,3000	29,4286	56,4810	47,5194	76,3214	35,5252
LR	30,7418	43,2133	39,9714	39,5833	47,0181	42,4310	47,7571	38,9424
SVM	**26,0939**	43,8857	32,9042	**16,7571**	53,5773	42,5929	47,2551	**27,0167**
NB	30,9963	42,7204	36,6650	32,0978	44,4188	34,1164	52,5429	34,1986
CART	31,3900	40,1138	29,0542	39,3833	40,7154	52,9689	48,0381	37,7877
RF	35,8737	40,2605	37,8917	28,0467	47,1972	44,3350	39,2579	33,1686
XGB	34,4830	42,2082	42,0583	36,5667	56,0688	42,7316	42,0010	39,9322
MLP	26,3647	**32,6717**	26,1400	24,7786	48,8528	**33,4903**	64,3214	36,8301
DBN	32,9707	34,7859	**22,1650**	19,2286	47,5580	40,8544	42,6071	33,1824
SAE	36,0962	37,6802	34,2600	23,6786	52,2396	35,0764	44,0238	40,3083
RNN	33,0159	37,4032	44,9229	26,6119	47,3606	35,6376	40,3143	36,5903
LSTM	33,8783	44,3280	43,7407	42,7250	45,8940	45,3289	50,4286	40,4576
GRU	35,1255	42,9967	37,6500	33,6900	**39,8872**	42,3257	**28,9762**	42,5434

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 41 – Tempo Médio de Recuperação (ART) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**18,9086**	**15,6808**	**17,1746**	16,0000	**15,1359**	**15,7629**	12,5667	**14,8899**
BAH	79,5644	121,8397	25,0076	17,8400	31,7313	47,2538	66,7614	75,8410
LR	78,1743	83,5508	64,5919	15,0400	35,2806	72,7485	69,0417	85,0642
SVM	127,6614	73,0758	69,3294	16,7727	23,4789	69,4506	237,8636	64,4154
NB	89,1895	**56,0132**	32,3111	13,0952	83,8174	97,1430	73,0250	87,2842
CART	90,9339	83,4962	40,9733	12,9630	93,4241	61,8436	38,9018	44,1707
RF	129,9493	71,1996	60,7722	10,6667	**17,5068**	51,5747	40,7857	78,5238
XGB	34,8078	143,3285	54,4206	14,4286	38,9102	**46,5137**	122,6500	82,4505
MLP	83,3577	80,3769	**17,1795**	52,5556	33,0864	107,9099	**11,0000**	124,3262
DBN	74,1430	134,8284	24,1763	**0,0000**	80,5151	54,0338	71,2250	118,1126
SAE	**32,3167**	110,8482	187,9444	52,6667	83,2300	114,8608	69,5000	74,3454
RNN	78,9853	99,2676	19,1846	59,2857	34,5222	95,0173	74,1250	75,5338
LSTM	74,0629	130,9088	45,9524	39,1667	108,8584	86,1912	37,7411	**43,0515**
GRU	111,4464	130,9605	113,1696	15,7500	52,0621	67,4228	90,6667	87,4836

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 42 – Tempo Médio de Recuperação (ART) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	38,5000	**38,6667**	38,6667	38,6667	**38,6667**	38,6667	38,6667	38,6667
BAH	31,9389	54,4205	26,7857	22,0000	61,9683	35,5344	43,3750	32,0208
LR	52,0417	53,2692	73,1250	13,2857	51,8958	50,6458	27,9583	31,6924
SVM	41,2569	52,6952	32,9583	8,5714	53,8576	31,7756	29,8333	30,2439
NB	**22,0972**	55,3154	28,2500	37,0000	**40,1250**	38,4896	40,8750	**23,8639**
CART	41,7444	60,4679	68,1667	15,8333	64,6979	52,5677	34,6667	24,9236
RF	43,2083	53,5385	29,5667	4,7778	51,1354	43,2292	24,3571	32,2924
XGB	65,8889	46,2722	77,3333	11,1250	63,2250	51,8242	**20,8125**	30,8021
MLP	40,2583	48,7538	14,6250	22,0000	61,9683	**26,3469**	29,2500	39,4167
DBN	38,0083	**43,6502**	19,0000	**0,0000**	47,9683	33,7635	55,7500	29,8646
SAE	24,3694	50,1714	**7,7857**	13,2500	61,9683	41,0844	43,3750	35,5625
RNN	64,0333	50,2218	66,7857	5,6667	48,0461	43,1313	32,5556	35,3158
LSTM	52,4167	44,8821	25,3333	7,6000	66,4167	66,2042	36,7500	32,6914
GRU	51,5764	45,1731	32,1000	2,5000	54,1146	28,4202	25,2500	23,8830

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 43 – *Sharpe Ratio* Anualizado (ASR) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,9893**	**1,0941**	1,0291	1,1077	1,1259	**1,1001**	1,4456	**1,1651**
BAH	0,7291	0,6957	0,7319	0,8311	0,8763	0,8704	0,7073	0,8974
LR	0,7624	0,7374	0,5889	1,3261	1,1013	0,7063	−0,4035	0,7614
SVM	0,6725	0,6522	0,2729	1,5386	1,1801	0,6609	**1,5048**	0,6772
NB	0,3473	**0,7438**	0,8855	1,8613	1,0315	0,7443	0,6483	0,6905
CART	0,6492	0,6048	0,6925	1,6422	0,6449	0,7196	1,3102	0,7845
RF	0,6917	0,5325	0,5255	**2,4813**	**1,4480**	0,7526*	1,0764	0,8603*
XGB	**0,8517**	0,5041	1,0176	1,5648	1,0052	0,5037	1,2154	0,7127
MLP	0,5760	0,7305	0,3605	0,4339	0,8852	0,5979	0,9935	0,0683
DBN	0,4397	0,1857	0,2896	0,0000	0,8269	0,7415	0,8837	0,3915
SAE	0,6784	0,5643	−0,0448	0,5060	0,8228	0,5226	0,4230	0,7292
RNN	0,5601	0,5288	**1,2449**	0,7564	0,9799	0,4620	0,0787	0,8371
LSTM	0,8340	0,4126	0,7564	0,4082	0,6556	0,3982	1,0720	0,7045
GRU	0,5899	0,3121	0,7179	1,2806	0,7290	0,4011	0,7392	0,5077

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 44 – *Sharpe Ratio* Anualizado (ASR) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,4810	−0,5119	−0,5119	−0,5119	−0,5119	−0,5119	−0,5119	−0,5119
BAH	−0,1087	−0,5772	−0,0906	0,4203	−0,5007	0,2048	−0,4536	−0,0465
LR	−0,2786	−0,5488	−1,1109	0,3256	−0,5332	−0,0918	−0,2318	0,0480
SVM	0,2968	−0,9223	−0,1024	1,7488	−1,1976	0,1586	0,2425	0,5992
NB	**1,1662**	−0,7244	0,4463	0,4429	**0,0449**	0,3130	−0,8447	0,1452
CART	−0,1425	−0,9358	−0,6368	1,2389	−0,5916	−0,0157	0,2445	0,4256
RF	0,0961	−1,2611	−0,2280	**3,5989**	−1,3644	−0,1600	−0,1435	**0,6930**
XGB	−0,1604	−0,1072	−0,7452	2,0524	−0,5242	−0,3079	**0,8717**	0,1531
MLP	−0,9261	−0,6893	**0,8996**	0,4203	−0,7850	**0,6755**	−0,8429	−0,5164
DBN	−0,6943	−**0,0274**	−0,0745	0,0000	−0,3308	0,3768	−1,6318	−0,0567
SAE	−0,1238	−0,3431	0,3245	0,7124	−0,5474	0,5585	−0,2888	−0,4582
RNN	0,1731	−0,3314	−0,6365	2,7810	−0,0156	0,5439	0,2652	0,2697
LSTM	−0,3453	−0,0327	0,1017	1,3760	−0,0622	−0,1573	−0,7231	−0,4774
GRU	0,0670	−0,1488	0,4185	3,1546	−0,0789	0,3960	−0,7661	0,5944

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 45 – Índice Calmar (CAR) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	1,2307	1,3287	1,2200	2,0912	**4,5607**	**2,0772**	1,0810	1,8088
BAH	0,7784	1,2603	1,0433	1,9741	2,8446	1,3151	0,6894	2,7726
LR	0,7176	1,4125	1,1623	0,4231	3,5689	1,4263	0,8824	2,9857
SVM	0,7168	**1,7395**	1,2030	2,4267	3,6729	1,4499	**1,3274**	2,6855
NB	**1,4414**	1,4272	0,9353	1,4933	**4,2678**	0,9894	1,0834	2,7296
CART	0,6254	1,4186	**1,2589**	**2,4694**	3,0957	1,6873	0,9695	2,6172
RF	1,0712	1,4294	1,1150	1,6606	3,8712	1,4740	1,2023	**3,2828**
XGB	0,8501	1,4850	1,0099	0,9126	3,8699	**1,7811**	0,2916	2,6130
MLP	0,4414	1,1665	0,5683	2,0908	1,4836	1,2534	0,3682	2,1787
DBN	0,4081	1,1769	1,1419	0,1705	2,6837	1,1384	0,1941	2,2691
SAE	0,5328	1,0494	1,0579	2,1322	2,7445	1,1852	0,1792	2,2318
RNN	0,6782	1,2486	1,0846	2,1109	2,8568	1,0690	0,7401	2,5117
LSTM	0,7667	1,2720	0,8220	2,2663	2,6737	1,1962	0,4905	2,3385
GRU	0,7279	1,1483	0,7470	1,8570	2,7326	1,0403	0,3548	2,5047

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 46 – Índice Calmar (CAR) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,7616	−0,9600	−0,6170	−0,6141	−0,6543	−0,4587	−0,6532	−0,5829
BAH	0,1050	0,4869	0,7366	1,4338	−0,4204	0,5413	0,4069	0,1472
LR	0,9513	1,1227	0,9163	0,8030	−0,2367	1,5929	−0,3249	1,0455
SVM	1,0877	0,9018	0,6194	**3,6573**	−0,3112	1,6190	0,8256	**1,5893**
NB	**10,5530**	1,0624	−0,0336	2,4448	0,1408	**2,9795**	0,7726	1,2618
CART	0,5671	0,9305	**1,8242**	0,2043	0,1510	0,5030	0,1296	0,4171
RF	0,2118	0,6080	0,0765	2,4824	0,1872	0,9675	1,3429	1,4233
XGB	0,8627	1,0174	0,3619	0,4918	0,2796	0,5622	0,9968	0,6278
MLP	0,2193	0,7142	1,6348	2,6636	−0,2267	0,9360	−0,1626	0,0683
DBN	−0,1365	0,5944	1,6317	1,7959	−0,1574	1,2172	0,2987	0,1864
SAE	0,0445	0,5350	1,5580	1,3164	−0,2666	0,9496	−0,0022	0,2526
RNN	0,8396	0,9011	0,6988	2,8044	0,1417	0,7288	1,1548	0,6445
LSTM	0,7657	0,3626	1,0956	1,0751	−0,4006	0,5778	0,0336	0,7778
GRU	1,2056	**1,1538**	0,2479	1,3974	**1,0333**	1,5063	**2,0603**	0,7817

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 47 – Índice Calmar (CAR) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	0,9778	1,1069	1,0211	1,0930	1,3118	**1,0979**	2,1907	1,5564
BAH	1,0571	1,1487	0,7692	1,0634	1,1617	1,0742	1,8253	1,5724
LR	1,0705	**1,2495**	0,5166	2,3717	1,4674	0,8960	−0,1830	1,4953
SVM	1,0261	1,1091	0,2877	3,3313	1,7807	0,9662	**5,3310**	1,4878
NB	0,4786	1,1252	1,1777	**4,4264**	1,2940	1,0293*	1,4559	1,5691
CART	0,8981	1,0287	0,7566	2,8944	0,8379	0,9626	3,7557	1,4465
RF	1,0322	0,9382	0,4631	4,4201	**1,9288**	0,8592	3,0604	**1,7497**
XGB	1,1041	1,0074	1,2639	2,6074	1,0324	0,5054	3,6048	1,2822
MLP	0,7614	1,1683	0,3536	0,4693	1,1023	0,8021	2,3964	0,9423
DBN	0,6373	0,4914	0,2518	0,0000	1,0622	0,9418	2,0990	1,2207
SAE	0,9919	1,0298	−0,0347	0,5281	1,1021	0,7441	1,5033	1,4754
RNN	0,7991	0,8695	**1,6740**	0,6774	1,1160	0,5235	0,9146	1,4225
LSTM	**1,2022**	0,7170	1,2585	0,4044	0,9143	0,4729	2,2325	1,3645
GRU	0,8725	0,5522	1,0930	1,2979	0,9066	0,5560	1,9113	1,1114

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 48 – Índice Calmar (CAR) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,6193	−0,6532	−0,6532	−0,6532	−0,6532	−0,6532	−0,6532	−0,6532
BAH	−0,1235	−0,7903	−0,1608	0,9863	5,0676	0,3923	−0,6610	−0,0468
LR	−0,0425	0,0837	−1,4320	0,5959	−0,5797	0,2650	−0,1364	0,6511
SVM	0,7565	−0,9324	−0,1707	6,4731	−0,5522	0,5626	0,4526	1,3521
NB	**3,2146**	−0,4916	1,0519	1,0648	1,2394	1,3134	−0,9653	0,8127
CART	−0,1292	−0,8765	−0,8448	2,8812	−0,4206	0,3377	0,4391	1,0582
RF	0,4918	−1,3970	−0,2281	14,2510	−1,4847	−0,0052	0,4475	**1,6143**
XGB	−0,2186	0,4969	−0,9817	7,3471	7,4636	−0,0929	**1,8027**	0,6556
MLP	−1,0440	−0,6850	**3,7280**	0,9863	5,0035	**2,8476**	−0,9499	−0,2873
DBN	−0,7672	**1,7885**	−0,1361	0,0000	5,4235	0,9290	−1,7319	0,2718
SAE	−0,1783	0,0580	0,5754	2,1019	5,1419	1,3654	−0,4075	−0,2817
RNN	1,4535	−0,0600	−0,9531	12,3634	**7,9951**	0,9574	1,2064	0,6091
LSTM	−0,4340	1,3153	0,8264	4,2638	5,3762	−0,0246	−0,9834	−0,3886
GRU	0,4070	0,3422	1,5407	**20,7712**	5,7038	2,3375	−1,0729	1,5485

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 49 – Máximo *Drawdown* (MDD) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,1832**	**0,1915**	**0,1922**	**0,1747**	**0,1831**	**0,1811**	**0,2035**	0,1799
BAH	0,4022	0,3875	0,4142	0,6273	0,3323	0,4947	0,5797	0,2501
LR	0,3260	0,2986	0,3048	0,5030	0,2474	0,3804	0,4218	0,1730
SVM	0,3268	0,2774	0,3264	0,5030	0,2606	0,4114	0,3850	0,1889
NB	0,3329	**0,2630**	0,3119	0,4491	**0,2197**	**0,3565**	**0,3670**	**0,1611**
CART	0,3171	0,2943	0,2966	0,4984	0,2520	0,3743	0,3834	0,2010
RF	**0,3024**	0,2929	0,3112	0,4350	0,2604	0,3762	0,3806	0,1769
XGB	0,3098	0,2913	0,3438	0,5235	0,2649	0,3928	0,5059	0,1935
MLP	0,3166	0,2841	0,2938	0,4172	0,2404	0,3896	0,5461	0,2024
DBN	0,3366	0,3249	**0,2786**	**0,2292**	0,2884	0,4391	0,5161	0,2003
SAE	0,3239	0,3172	0,3306	0,4340	0,2916	0,4601	0,4736	0,2274
RNN	0,3364	0,3002	0,3135	0,4398	0,3066	0,4247	0,4488	0,2187
LSTM	0,3406	0,3195	0,3272	0,4256	0,3080	0,4137	0,4779	0,2133
GRU	0,3332	0,3336	0,3310	0,4673	0,3000	0,4117	0,5075	0,2007

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 50 – Máximo *Drawdown* (MDD) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	0,4605	0,4621	0,4545	0,4594	0,4628	0,4502	0,4682	0,4483
BAH	0,5459	0,5532	0,4388	0,5094	0,5132	0,5989	0,6319	0,3992
LR	0,3821	0,3801	0,3394	0,3356	0,3710	0,3814	0,4985	0,3058
SVM	0,3718	0,3881	0,3342	0,3330	0,3741	0,3846	0,4346	0,2847
NB	0,3815	**0,3296**	0,3367	0,3907	0,3202	**0,2932**	0,4667	**0,2801**
CART	0,4110	0,4026	0,3017	0,4510	0,3584	0,4336	0,5162	0,3181
RF	0,4404	0,4192	0,3506	0,3408	0,3704	0,4395	0,4829	0,3027
XGB	0,3696	0,4012	0,3720	0,4266	0,3847	0,4373	0,4798	0,3158
MLP	0,3405	0,4032	0,2908	0,3957	**0,2874**	0,3465	0,5338	0,3109
DBN	0,3476	0,4420	0,3019	**0,2771**	0,3594	0,3894	0,3993	0,2991
SAE	0,3858	0,4397	0,3665	0,3780	0,3896	0,4406	0,4109	0,3154
RNN	0,3630	0,3798	0,3235	0,3775	0,3778	0,3946	0,3838	0,3387
LSTM	0,3809	0,4455	0,3436	0,3512	0,3886	0,4360	0,4134	0,3238
GRU	**0,3213**	0,3574	**0,2184**	0,3652	0,3336	0,3394	**0,3070**	0,2972

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 51 – Máximo *Drawdown* (MDD) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	0,2024	**0,1993**	0,2035	0,2035	**0,1883**	**0,2015**	0,1461	**0,1746**
BAH	0,3325	0,4728	0,4086	0,1935	0,3662	0,4184	0,3628	0,3755
LR	0,2218	0,3175	0,3676	0,1071	0,2700	0,3441	0,3129	0,2995
SVM	0,2898	0,3648	0,4058	0,0778	**0,2352**	0,3754	0,2486	0,3190
NB	0,2473	**0,3062**	0,2062	0,0693	0,2593	**0,3104**	0,2390	0,2830
CART	0,2455	0,3574	0,3352	0,0974	0,2986	0,3634	0,1943	0,2850
RF	0,2734	0,3731	0,3506	0,0927	0,2724	0,3468	0,2344	0,2721
XGB	0,2805	0,4098	0,2024	0,1130	0,3224	0,3566	0,2046	0,3220
MLP	0,3064	0,3731	0,3088	0,2018	0,3379	0,4188	**0,0440**	0,3629
DBN	0,2507	0,3710	0,3524	**0,0000**	0,3396	0,4266	0,1333	**0,2492**
SAE	**0,1696**	0,3682	0,2493	0,1521	0,3387	0,3938	0,2928	0,2642
RNN	0,3183	0,4402	0,2236	0,1935	0,2753	0,3723	0,2836	0,2907
LSTM	0,3042	0,4211	0,2829	0,1877	0,3232	0,3848	0,1971	0,3025
GRU	0,2901	0,4129	**0,1483**	0,1960	0,2796	0,3680	0,2320	0,3247

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 52 – Máximo *Drawdown* (MDD) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	0,4628	0,4682	0,4682	0,4682	0,4682	0,4682	0,4682	0,4682
BAH	0,5452	0,6377	0,4378	0,2326	0,5316	0,5709	0,5844	0,4222
LR	0,3984	0,4334	0,3652	0,1579	0,3522	0,3570	0,3212	0,3220
SVM	0,3363	0,4264	0,2994	0,0939	0,4044	0,3755	0,3196	0,2845
NB	0,2313	0,3930	0,2210	0,1648	**0,2863**	0,3389	0,4975	0,2713
CART	0,3501	0,4849	0,3425	0,1493	0,2991	0,4158	0,3979	0,3121
RF	0,3475	0,5074	0,2607	0,0939	0,4318	0,4337	0,3917	0,3084
XGB	0,3951	0,4139	0,2454	0,0872	0,4331	0,4384	0,3618	0,3012
MLP	0,4431	0,4790	**0,0646**	0,2326	0,4564	0,3736	0,2917	0,2561
DBN	0,3247	**0,3234**	0,1445	**0,0000**	0,4612	0,4292	0,2987	**0,2301**
SAE	**0,2019**	0,3767	0,2147	0,1663	0,4622	0,3842	0,4454	0,2513
RNN	0,2875	0,4196	0,3059	0,0780	0,3698	**0,2280**	**0,2186**	0,3183
LSTM	0,4044	0,4427	0,3320	0,0568	0,3676	0,3647	0,4064	0,3141
GRU	0,4094	0,3399	0,2088	0,0483	0,3958	0,3058	0,3929	0,2686

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. Para esta métrica, **menores** valores indicam melhor desempenho. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 53 – *Sharpe Ratio* Modificado (MSR) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**1,0198**	**1,0501**	**1,0482**	**1,1176**	**1,3908**	**1,2429**	**1,0365**	**1,1762**
BAH	0,5006	0,3853	−0,3940	−0,4235	0,6172	0,2260	0,5621	−1,1623
LR	0,1742	0,0095	−1,2445	0,0351	−0,5496	−0,2037	0,4656	−1,5275
SVM	0,3803	0,0236	−0,7706	−0,7144	−0,3755	−0,3471	−0,8094	−1,7858
NB	0,3125	0,2867*	−**0,0653**	−0,1095	−1,9477	−0,0001	0,2681	−2,1553
CART	0,4182*	0,0630	−0,3346	−0,8626	0,3027	0,0441	0,0269	−2,2118
RF	0,1774	−0,3477	−0,7845	−0,3328	−0,6070	−0,4606	0,0288	−2,9478
XGB	0,3864	−0,6781	−1,2433	**0,2150**	−0,0837	−0,8677	0,1401	−1,2532
MLP	0,2990	0,2381	−0,5173	−0,3527	0,0896	0,0828*	0,4237	−1,2737
DBN	0,2325	0,2414	−0,2979	0,1428	0,1039	0,0536	−1,1248	−**0,4725**
SAE	0,3298	0,2241	−0,4819	−0,4197	0,5355*	0,0581	**0,6639**	−0,9293
RNN	0,3162	0,1344	−1,2248	−0,2884	−0,2071	0,0696	0,4841	−0,8003
LSTM	0,4152	0,1149	−0,4909	−0,2698	0,1067	−0,0176	0,1707	−1,0397
GRU	−0,0160	0,1175	−0,3586	−0,2013	0,1606	0,0628	0,1373	−0,5010

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 54 – *Sharpe Ratio* Modificado (MSR) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,4804	−0,2576	−0,4801	−0,4801	−0,5100	−0,4397	−0,5172	−0,4587
BAH	−0,0345	−0,0543	0,4121	**0,6367**	−0,3112	0,0690	0,0928	−0,2951
LR	0,0492	0,1753	0,4536	0,3179	−0,0285	−0,0289	0,1130	−0,3697
SVM	−0,1451	−0,2062	0,2184	−0,6966	−0,2788	0,0268	0,0980	−0,6224
NB	−2,4693	−0,1597	0,1305	0,5824	0,0446	−0,9656	−0,5247	−0,9712
CART	−0,2897	−0,2100	0,4355	0,1987	0,0540	0,0951	−0,0301	−0,4772
RF	−0,2398	−0,1742	−0,0584	0,0218	−0,2398	−0,1235	−1,7100	−1,2724
XGB	0,3783	−0,0855	−0,0020	0,0894	−0,1361	−0,2145	−0,5587	−0,3112
MLP	0,0559	**0,2598**	0,3667	−0,4075	**0,4637**	−0,1553	**0,1776**	0,1156
DBN	**0,7734**	−0,0867	0,5660	−0,7610	0,1883	−0,1662	0,1449	−0,3205
SAE	−0,0680	−0,0530	0,0793	**0,5920**	−0,1042	−0,0143	−0,0190	**0,4429**
RNN	−0,1177	−0,3941	0,3565	0,0086	0,0289	**0,4367**	−1,4260	−0,2984
LSTM	−0,1859	0,0014	−1,1772	0,3618	0,1837	0,1221	−2,1637	−0,4214
GRU	0,0804	−0,7009	**1,0220**	0,2468	−0,4225	−0,2501	−1,3227	−0,2297

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 55 – *Sharpe Ratio* Modificado (MSR) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,9389**	**1,0334**	**0,9802**	1,0477	**1,0625**	**1,0410**	**1,2614**	1,0958
BAH	0,5620	0,6001	0,5060	0,8208	0,7826	0,7421	0,7745	0,7810
LR	**0,6569**	0,3555	0,2634	0,8385	0,4203	0,4852	−0,3662	0,5860
SVM	0,2551	0,4538	0,1873	1,0654	0,4092	0,4858	−1,5386	0,4324
NB	0,2561	0,2087	0,5914	−1,1106	0,0800	0,4538	0,6213	0,2690
CART	0,5304	0,4983	0,6180	1,0522	0,5258	0,5190	0,8371	0,3386
RF	0,3725	0,4746	0,4054	−3,5229	0,0700	0,6312	0,7221	0,5559
XGB	0,3235	0,3041	0,6644	**1,1570**	0,5078	0,4454	0,7286	0,5249
MLP	0,4751	**0,6765**	0,2779	0,4350	0,7747*	0,5837	**0,9896**	**8,9419**
DBN	0,3857	0,4188	0,2788	0,0000	0,6902	0,6799*	0,2979	0,9224
SAE	0,4819	0,5121	−0,0467	0,4875	0,6946	0,3922	0,5216	0,7133
RNN	0,3771	0,4028	**0,7152**	0,7042	0,6542	0,2756	0,0705	0,6713
LSTM	0,5467	0,3870	0,4914	0,4067	0,5123	0,3315	0,7269	0,5484
GRU	0,4262	0,2818	0,2276	0,8849	0,3890	0,3383	0,7591	0,8459

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 56 – *Sharpe Ratio* Modificado (MSR) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,4864	−0,5172	−0,5172	−0,5172	−0,5172	−0,5172	−0,5172	−0,5172
BAH	−0,1066	−0,4643	−0,0901	0,4113	−1,9328	0,0288	−0,4418	−0,0504
LR	−0,3561	−0,5230	−0,7030	0,3245	−0,2955	0,0137	−0,0418	0,0518
SVM	0,1624	0,1281	−0,1018	0,3763	−1,1428	0,0702	0,2178	0,4740
NB	0,6803	**4,5063**	0,3945	0,4546	−1,5201	−0,2781	−0,6262	−0,2368
CART	−0,1190	0,7594	−0,4769	0,8300	−0,2308	−0,1609	0,2442	0,2473
RF	0,0909	0,8717	−0,0843	−10,6665	**−0,0991**	−0,5222	−0,1681	−0,0359
XGB	−0,1119	0,2001	−0,4880	−0,1847	−1,6329	**0,2151**	**0,6653**	0,0739
MLP	0,5550	0,4892	−2,5889	0,4113	−0,6485	−2,2033	−0,7274	**2,7997**
DBN	**2,1759**	−1,1488	−0,0739	0,0000	−1,5909	0,0446	0,0791	−0,1503
SAE	−0,1052	0,0983	0,2750	0,6765	−1,2622	0,0482	−0,2799	0,6450
RNN	0,1638	−0,2779	**2,0551**	−7,4299	−1,5611	0,2106	0,2706	0,0722
LSTM	−0,2847	−0,9021	0,1697	**1,0272**	−1,2880	−0,1671	−0,5624	0,3031
GRU	0,0372	−0,1378	0,3427	−16,8628	−1,4463	−1,3794	−0,5896	−0,0856

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 57 – Índice Sortino (SOR) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,0690**	**0,0715**	**0,0708**	**0,0839**	**0,0865**	**0,0827**	**0,0695**	0,0798
BAH	0,0442	0,0554	0,0346	0,0647	0,0767	0,0592	0,0467	**0,1014**
LR	0,0255	0,0428	0,0336	0,0061	0,0728	0,0396	0,0387	0,0742
SVM	0,0297	0,0550*	**0,0396**	0,0543	0,0741*	0,0444	**0,0589**	0,0777
NB	0,0354	0,0408	0,0343	0,0328	0,0734	0,0317	0,0457	0,0680
CART	0,0277	0,0462	0,0348	0,0556	0,0599	0,0488	0,0421	0,0668
RF	0,0408*	0,0459	0,0357	0,0424	0,0711	0,0419	0,0471	0,0770
XGB	0,0314	0,0449	0,0329	0,0265	0,0679	0,0473	0,0120	0,0699
MLP	0,0248	0,0415	0,0186	0,0676	0,0523	0,0489*	0,0242	0,0708
DBN	0,0204	0,0450	0,0376	0,0091	0,0604	0,0444	0,0127	0,0771
SAE	0,0254	0,0374	0,0328	0,0660	0,0659	0,0476	0,0120	0,0807
RNN	0,0289	0,0372	0,0348	0,0724	0,0605	0,0403	0,0408	**0,0810**
LSTM	0,0344	0,0428	0,0284	**0,0767**	0,0580	0,0428	0,0248	0,0687
GRU	0,0311	0,0372	0,0206	0,0606	0,0622	0,0362	0,0213	0,0747

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 58 – Índice Sortino (SOR) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,0349	−0,0384	−0,0279	−0,0278	−0,0296	−0,0245	−0,0299	−0,0265
BAH	−0,0013	0,0025	0,0299	0,0524	−0,0232	0,0142	0,0122	0,0009
LR	0,0150	**0,0129**	0,0271	0,0175	−0,0185	0,0334	−0,0137	0,0130
SVM	0,0198	0,0086	0,0129	**0,0885**	−0,0164	0,0340	0,0193	**0,0298**
NB	**0,0244**	0,0119	−0,0106	0,0724	−0,0078	**0,0449**	0,0116	0,0039
CART	0,0042	0,0116	0,0393	0,0068	−0,0057	0,0069	0,0014	0,0022
RF	−0,0041	0,0021	−0,0082	0,0595	−0,0117	0,0206	0,0263	0,0170
XGB	0,0086	0,0097	0,0007	0,0152	−0,0066	0,0105	0,0183	0,0047
MLP	−0,0053	0,0086	**0,0448**	0,0531	−0,0122	0,0116	−0,0040	−0,0026
DBN	−0,0069	0,0026	0,0425	0,0233	−0,0126	0,0207	0,0103	−0,0007
SAE	−0,0076	0,0019	0,0353	0,0449	−0,0185	0,0144	0,0013	0,0022
RNN	0,0163	0,0094	0,0217	0,0442	−0,0061	0,0197	0,0202	0,0101
LSTM	0,0062	−0,0015	0,0187	0,0144	−0,0183	0,0091	−0,0020	0,0101
GRU	0,0150	0,0082	0,0101	0,0274	**−0,0027**	0,0270	**0,0344**	0,0130

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 59 – Índice Sortino (SOR) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,0626**	**0,0693**	0,0654	0,0703	0,0716	**0,0698**	**0,0906**	**0,0743**
BAH	0,0492	0,0475	0,0483	0,0565	0,0574	0,0579	0,0474	0,0601
LR	0,0393	0,0405	0,0295	0,0727	0,0570	0,0355	−0,0181	0,0436
SVM	0,0344	0,0381	0,0151	0,0760	0,0624	0,0354	**0,0789**	0,0404
NB	0,0137	0,0363	0,0450	0,0905	0,0455	0,0341	0,0390	0,0395
CART	0,0350	0,0311	0,0364	0,0781	0,0329	0,0360	0,0606	0,0444
RF	0,0351	0,0287	0,0243	**0,1311**	**0,0768**	0,0375	0,0482	0,0470
XGB	0,0423	0,0293	0,0541	0,0832	0,0518	0,0237	0,0590	0,0400
MLP	0,0354	**0,0497**	0,0122	0,0266	0,0536	0,0367	0,0578	0,0309
DBN	0,0311	0,0173	0,0215	0,0000	0,0482	0,0484*	0,0343	0,0306
SAE	0,0406	0,0394	−0,0019	0,0217	0,0498	0,0290	0,0364	0,0460
RNN	0,0319	0,0353	**0,0710**	0,0360	0,0446	0,0243	0,0013	0,0509*
LSTM	0,0473*	0,0271	0,0361	0,0214	0,0346	0,0227	0,0450	0,0409
GRU	0,0353	0,0176	0,0272	0,0685	0,0351	0,0238	0,0441	0,0340

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 60 – Índice Sortino (SOR) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	−0,0281	−0,0299	−0,0299	−0,0299	−0,0299	−0,0299	−0,0299	−0,0299
BAH	−0,0070	−0,0369	−0,0064	0,0255	0,2340	0,0121	−0,0296	−0,0037
LR	−0,0135	−0,0175	−0,0523	0,0135	−0,0196	0,0016	−0,0016	0,0084
SVM	0,0169	−0,0331	−0,0052	0,0835	−0,0409	0,0107	0,0116	0,0313
NB	**0,0676**	−0,0191	0,0197	0,0213	0,0181	0,0181	−0,0362	0,0116
CART	−0,0062	−0,0335	−0,0253	0,0700	−0,0203	0,0036	0,0130	0,0231
RF	0,0066	−0,0504	−0,0048	**0,2234**	−0,0580	−0,0017	0,0014	**0,0355**
XGB	−0,0089	0,0047	−0,0325	0,1184	0,2257	−0,0120	**0,0506**	0,0072
MLP	−0,0370	−0,0348	**0,0285**	0,0255	0,2396	**0,0390**	−0,0481	−0,0038
DBN	−0,0221	**0,0058**	−0,0055	0,0000	0,2439	0,0330	−0,0707	−0,0070
SAE	−0,0044	−0,0095	0,0195	0,0424	0,2450	0,0317	−0,0192	−0,0118
RNN	0,0198	−0,0137	−0,0061	0,1207	0,2303	0,0216	0,0243	0,0112
LSTM	−0,0150	−0,0013	0,0103	0,0562	0,2387	−0,0088	−0,0289	−0,0141
GRU	0,0084	−0,0055	0,0096	0,1598	**0,2615**	0,0181	−0,0306	0,0291

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 61 – Taxa de Acerto (WR) no cenário InS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,5340**	**0,5351**	**0,5345**	**0,5394**	**0,5349**	**0,5354**	0,5341	0,5359
BAH	0,4947	0,5003	0,4875	0,4417	0,5062	0,4866	0,4808	0,5389
LR	0,5016	0,5059	0,4940	0,4474	0,5240	0,5011	0,4977	0,5445
SVM	0,5021	**0,5144**	0,4992	0,4408	0,5224	0,5014	0,5087	0,5406
NB	0,5057	0,5143	0,5022	**0,4696**	0,5233	0,5004	0,5033	**0,5696**
CART	0,5050	0,5091	0,4964	0,4512	0,5168	**0,5051**	0,4797	0,5393
RF	**0,5089**	0,5127	**0,5035**	0,4496	**0,5298**	0,5026	0,5127	0,5480
XGB	0,5024	0,5095	0,4974	0,4214	0,5151	0,4960	0,4845	0,5456
MLP	0,4238	0,4279	0,3852	0,4435	0,4149	0,4548	0,4705	0,4947
DBN	0,4445	0,4679	0,4342	0,2739	0,4728	0,4647	**0,5398**	0,4684
SAE	0,4504	0,4560	0,4833	0,4409	0,4730	0,4686	0,3959	0,5151
RNN	0,4973	0,5004	0,4863	0,4448	0,5126	0,4889	0,4917	0,5384
LSTM	0,5013	0,5060	0,4814	0,4220	0,5117	0,4864	0,4917	0,5384
GRU	0,4989	0,4984	0,4754	0,4298	0,5086	0,4885	0,4891	0,5431

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 62 – Taxa de Acerto (WR) no cenário InS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	0,4888	0,4896	0,4899	0,4916	**0,4920**	0,4935	0,4917	0,4949
BAH	0,4999	0,4793	0,5019	0,5150	0,4786	0,4800	0,4828	0,5165
LR	0,4981	0,4918	0,5149	0,5077	0,4717	0,4943	0,5000	0,5336
SVM	0,5029	0,4925	0,5033	**0,5527**	0,4672	0,4896	0,5173	**0,5472**
NB	0,5128	0,4850	0,4727	0,5344	0,4863	**0,5138**	0,5196	0,5316
CART	0,4928	0,4886	0,5089	0,4983	0,4721	0,4719	0,4980	0,5195
RF	0,4946	0,4848	0,4946	0,5406	0,4709	0,4861	0,5238	0,5355
XGB	0,5076	0,4877	0,4975	0,5078	0,4768	0,4821	0,5132	0,5208
MLP	0,3822	0,4090	0,4126	0,5421	0,3706	0,3760	0,4445	0,4665
DBN	0,4091	0,4326	0,4700	0,4395	0,4174	0,4264	0,3321	0,4674
SAE	0,4322	0,4319	0,5225	0,4206	0,4414	0,4246	0,3466	0,4822
RNN	**0,5184**	**0,4935**	0,5091	0,5416	0,4694	0,4103	0,5095	0,5319
LSTM	0,4982	0,4867	**0,5339**	0,5145	**0,4868**	0,4898	**0,5594**	0,5290
GRU	0,5150	0,4907	0,4001	0,5088	0,4862	0,4873	0,5350	0,5304

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 63 – Taxa de Acerto (WR) no cenário OutS-TraD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,5305**	**0,5338**	0,5318	0,5346	0,5348	**0,5343**	0,5411	**0,5347**
BAH	0,4911	0,4881	0,4902	0,4990	0,5257	0,5071	0,5127	0,4990
LR	0,5024	**0,5021**	0,5040	0,5213	0,5343	0,5072	0,5050	**0,5102**
SVM	0,5004	0,4973	0,4944	0,5408	0,5372	0,5170	**0,5715**	0,5027
NB	0,4888	0,4954	0,4953	0,5616	0,5506	0,5156	0,4961	0,4918
CART	0,4877	0,4954	0,5026	0,5522	0,5181	0,5158	0,5519	0,4972
RF	0,4961	0,4948	0,4967	**0,5854**	**0,5507**	**0,5179**	0,5700	0,4988
XGB	**0,5153**	0,4926	0,5014	0,5203	0,5400	0,5135	0,5572	0,5058
MLP	0,4782	0,4499	**0,5554**	0,4916	0,5229	0,5023	0,2807	0,4633
DBN	0,3990	0,3896	0,3233	0,0000	0,5254	0,5020	0,5436	0,3655
SAE	0,3286	0,4434	0,3084	0,4870	0,5242	0,5051	0,5095	0,4225
RNN	0,4865	0,4772	0,4919	0,5042	0,5325	0,5145	0,5037	0,4980
LSTM	0,4947	0,4845	0,5177	0,4814	0,5286	0,4981	0,5206	0,4942
GRU	0,4881	0,4796	0,5093	0,5165	0,5326	0,5028	0,5062	0,4907

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.

Tabela 64 – Taxa de Acerto (WR) no cenário OutS-TesD

Modelo	**BM**	**CC**	**CNC**	**COM**	**FIN**	**GCS**	**OGB**	**UTI**
Index	**0,4923**	0,4917	0,4917	0,4917	0,4917	0,4917	0,4917	0,4917
BAH	0,4762	0,4655	0,4605	0,5546	0,4580	0,5076	0,4708	0,4812
LR	0,4593	0,4671	0,3994	0,5600	0,4641	0,5008	0,4657	0,4884
SVM	0,4669	0,4706	0,4350	0,5882	0,4327	0,4936	0,4804	0,4937
NB	0,4775	0,4282	0,4509	0,5185	**0,5012**	0,5068	0,4960	0,4937
CART	0,4614	0,4515	0,4386	0,5357	0,4561	0,5021	**0,5097**	0,4992
RF	0,4761	0,4540	0,3785	0,6667	0,4391	0,5073	0,4945	0,5248
XGB	0,4471	0,4912	0,4128	0,5846	0,4724	0,4673	0,5033	0,5001
MLP	0,4159	0,3846	**0,5935**	0,5546	0,4323	**0,5567**	0,2095	0,4185
DBN	0,3211	0,3783	0,2000	0,0000	0,4634	0,5104	0,4041	0,3610
SAE	0,3220	0,4227	0,2632	0,5455	0,4298	0,5050	0,4583	0,3856
RNN	0,4592	0,4769	0,4272	0,6538	0,4824	0,4840	0,4718	0,4960
LSTM	0,4639	**0,5026**	0,4345	0,5682	0,4731	0,4965	0,4313	0,4719
GRU	**0,4825**	0,4736	0,5576	**0,7647**	0,4447	0,5191	0,4606	**0,5263**

Nota: Valores em **negrito** indicam o melhor desempenho na coluna. O asterisco (*) sinaliza o melhor algoritmo quando os *benchmarks* superam todos os modelos de ML.