



FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

Aplicação de Técnicas de Machine
Learning na Optimização de Retornos
Financeiros no Mercado de Forex

Autor: Misael Eusébio Joaquim

Maputo, Junho de 2026



FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

Aplicação de Técnicas de Machine
Learning na Optimização de Retornos
Financeiros no Mercado de Forex

Autor: Misael Eusébio Joaquim

Supervisor: Prof. Doutor Cachimo Assane

Maputo, Junho de 2026

Declaração de Honra

Declaro por minha honra que este trabalho é resultado da minha investigação e que o processo foi concebido para ser submetido apenas para obtenção do grau de Licenciado em Estatística na Faculdade de Ciências da Universidade Eduardo Mondlane.

Maputo, Junho de 2026

(Misael Eusébio Joaquim)

Dedicatória

Dedico este trabalho a minha Mãe, Hortência Escola.

Agradecimentos

Agradeço ao meu supervisor, Prof. Doutor Cachimo Assane, pela paciência, pelo constante apoio e pela disponibilidade com que me guiou durante todo o processo de elaboração deste trabalho.

Ao corpo docente do Departamento de Matemática e Informática da Faculdade de Ciências, expresso a minha gratidão pela excelência no ensino e pelo rigor profissional. O conhecimento partilhado por cada um foi o alicerce indispensável sobre o qual edifiquei a minha trajectória académica.

À minha mãe, Hortência Escola, e à minha irmã, Páscoa, pelo apoio incondicional ao longo do meu percurso académico.

Por fim, a todos aqueles que, de forma directa ou indirecta, contribuíram para a realização deste trabalho, expresso o meu mais sincero agradecimento.

*"We're basically trying to predict the future,
and that's extremely hard."*

(Jim Simons)

Resumo

Este estudo investiga a aplicação de técnicas de machine learning na otimização de retornos financeiros no mercado de Forex. O estudo foi realizado com dados extraídos do metatrader 5, referentes ao período de 9 de Setembro de 2010 a 31 de Janeiro de 2025. Foram considerados modelos predictivos supervisionados, incluindo regressão logística penalizada via L1, árvores de decisão e redes neurais artificiais, avaliados por meio de validação cruzada temporal com janela expansiva e reamostragem bootstrap por blocos, assegurando a robustez e estabilidade dos resultados fora da amostra (30% da base de dados), com otimização de hiperparâmetros via grid search com validação cruzada temporal do tipo k-fold. Os resultados indicam que a regressão logística penalizada apresentou maior consistência e estabilidade predictiva, superando estatisticamente a árvore de decisão e exibindo desempenho comparável ao da rede neural artificial, com acurácia fora da amostra em torno de 60%. A análise de estabilidade temporal evidenciou que a expansão progressiva da amostra de treino pode melhorar o desempenho dos algoritmos. Do ponto de vista económico, as previsões foram incorporadas em estratégias de negociação com regras explícitas de gestão de risco, incluindo a utilização do Value at Risk como mecanismo de stop-loss. Essa integração permitiu converter vantagens estatísticas moderadas em melhorias significativas na relação risco-retorno, reflectidas no aumento dos índices de Sharpe e Sortino e na redução das perdas acumuladas. Os resultados sugerem que a utilidade prática dos modelos de machine learning não depende exclusivamente da complexidade algorítmica, mas da combinação entre parcimónia estatística, validação temporal rigorosa e controlo eficaz do risco. Em síntese, o estudo contribui para a literatura empírica ao demonstrar como técnicas de machine learning podem apoiar a tomada de decisão no mercado Forex quando aplicadas com rigor metodológico e interpretação cautelosa, mesmo em ambientes caracterizados por elevada incerteza e instabilidade.

Palavras-chave: Bootstrap por blocos; Forex; Gestão de risco; Machine learning; Negociação; Redes neurais; Regressão logística penalizada; Validação cruzada com janela expansiva.

Abstract

This work investigates the application of machine learning techniques in optimizing financial returns within the Forex market, combining rigorous statistical analysis with an economic evaluation of forecasts. The study was conducted using data extracted from MetaTrader 5, covering the period from September 9, 2010, to January 31, 2025. Supervised predictive models were considered, including L1-penalized logistic regression, decision trees, and artificial neural networks, evaluated using walk-forward temporal validation with an expanding window and block bootstrap resampling, ensuring the robustness and stability of out-of-sample results (30% of the dataset), with hyperparameters optimized via grid search with time series k-fold cross-validation. The results indicate that penalized logistic regression showed greater consistency and predictive stability, statistically outperforming the decision tree and exhibiting performance comparable to the artificial neural network, with out-of-sample accuracies around 60%. Temporal stability analysis revealed that the progressive expansion of the training sample can improve algorithmic accuracy. From an economic perspective, forecasts were incorporated into trading strategies with explicit risk management rules, including the use of Value at Risk (VaR) as a stop-loss mechanism. This integration allowed for the conversion of moderate statistical advantages into significant improvements in the risk-return ratio, reflected in increased Sharpe and Sortino ratios and reduced draw-downs. The findings suggest that the practical utility of machine learning models depends not exclusively on algorithmic complexity, but on the combination of statistical parsimony, rigorous temporal validation, and effective risk control. In summary, the study contributes to the empirical literature by demonstrating that machine learning techniques can support decision-making in the Forex market when applied with methodological rigor and cautious interpretation, even in environments characterized by high uncertainty and instability.

Keywords: Block bootstrap; Forex; Machine learning; Neural networks; Penalized logistic regression; Risk management; Trading; Walk-forward validation.

Lista de Abreviaturas

ADF	<i>Augmented Dickey-Fuller</i>
ADX	<i>Average Directional Index</i>
API	<i>Application Programming Interface</i>
ATR	<i>Average True Range</i>
AUC	<i>Area Under the Curve</i>
CAPM	<i>Capital Asset Pricing Model</i>
CART	<i>Classification and Regression Trees</i>
CFD	<i>Contract for Difference</i>
CMO	<i>Chande Momentum Oscillator</i>
CVaR	<i>Conditional Value at Risk</i>
DMI	<i>Directional Movement Index</i>
DXY	Índice do Dólar Norte-Americano
EDA	<i>Exploratory Data Analysis</i>
EUR/USD	Par cambial Euro / Dólar Americano
FDC	<i>Fluxo de Dinheiro de Chaikin</i>
Forex	<i>Foreign Exchange Market</i>
FTSE100	<i>Financial Times Stock Exchange 100 Index</i>
GA	<i>Genetic Algorithm</i>
GBP/USD	Par cambial Libra Esterlina / Dólar Americano
GRU	<i>Gated Recurrent Unit</i>
IFR	Índice de Força Relativa
KNN	<i>K-Nearest Neighbors</i>
LASSO	<i>Least Absolute Shrinkage and Selection Operator</i>
LSTM	<i>Long Short-Term Memory</i>
MAR	<i>Minimum Acceptable Return</i>
MLP	<i>Multilayer Perceptron</i>
MQL5	<i>MetaQuotes Language 5</i>
OHLC	<i>Open, High, Low, Close</i>
RBF	<i>Radial Basis Function</i>
ROC	<i>Receiver Operating Characteristic</i>
ROI	<i>Return on Investment</i>
RNA	Rede Neural Artificial
S&P500	<i>Standard & Poor's 500 Index</i>
STOXX50	<i>EURO STOXX 50 Index</i>

SVM	<i>Support Vector Machine</i>
USD/JPY	Par cambial Dólar Americano / Iene Japonês
VaR	<i>Value at Risk</i>
WTI	<i>West Texas Intermediate</i>
WOE	<i>Weight of Evidence</i>

Índice

1	INTRODUÇÃO	1
1.1	Problematização	2
1.2	Objectivos	3
1.2.1	Geral	3
1.2.2	Específicos	3
1.3	Justificação	3
1.4	Estrutura do trabalho	4
2	Revisão da Literatura	5
2.1	Conceitos básicos	5
2.2	Mercado de Forex	6
2.2.1	Breve histórico	6
2.2.2	Principais participantes	7
2.2.3	Mecanismos de negociação de moedas	7
2.2.4	Moedas negociadas e principais pares	8
2.2.5	Processos de negociação e principais tipos de ordem	8
2.2.6	Instrumentos de negociação de moedas	9
2.2.7	Análise fundamental e análise técnica	11
2.2.8	Relações intermercados	12
2.3	Hipóteses do passeio aleatório e do mercado eficiente	14
2.4	Indicadores técnicos	14
2.5	Medidas estatísticas	29
2.5.1	Medidas de forma da distribuição	29
2.5.2	Medidas de variabilidade	31
2.5.3	Medidas de dependência	32
2.6	Séries temporais	34
2.6.1	Operações sobre séries temporais	34
2.6.2	Propriedades de séries temporais	35
2.7	Estimadores de volatilidade	37
2.8	Medidas de risco e desempenho financeiro	40
2.9	Machine learning	47
2.9.1	Aprendizagem supervisionada	47

2.9.2	Etapas de abordagem de machine learning	47
2.10	Peso da evidência	71
2.11	MetaTrader5	72
2.12	Trabalhos relacionados	73
3	Material e Metodos	75
3.1	Material	75
3.2	Métodos	78
3.2.1	Análise exploratória de dados	78
3.2.2	Divisão do conjunto de dados	78
3.2.3	Pré-processamento de dados	78
3.2.4	Critério de selecção de variáveis	79
3.2.5	Algoritmos de machine learning	79
3.2.6	Medidas de avaliação de desempenho predictivo dos algoritmos de clas- sificação	81
3.2.7	Melhoramento dos modelos	82
3.2.8	Critérios de selecção de modelos para as estratégias de negociação . .	86
3.2.9	Estratégia de negociação	87
3.2.10	Análise de risco e desempenho financeiro	88
4	Resultados e Discussão	89
4.1	Análise exploratória	89
4.2	Selecção de variáveis	95
4.3	Resultados de desempenho geral dos algoritmos	97
4.3.1	Análise do desempenho geral dos algoritmos	97
4.3.2	Matriz de Confusão	98
4.3.3	Curvas ROC	99
4.3.4	Comparação entre os algoritmos	101
4.4	Avaliação do desempenho dos algoritmos utilizando bootstrap por blocos moveis	103
4.5	Avaliação do desempenho dos algoritmos utilizando validação cruzada com janela expansiva	104
4.6	Aplicação das estratégias de negociação e análise de risco financeiro	111
4.6.1	Resultados da estratégia de negociação no par GBP/USD	111
4.6.2	Resultados da estratégia de negociação no par USD/JPY	114
4.6.3	Resultados da estratégia de negociação no par EUR/USD	117
4.7	Discussão dos Resultados	120
5	Conclusões e Recomendações	123
5.1	Recomendações	123
	Referências	125

Lista de Figuras

2.1	Volume médio diário negociado do mercado Forex por instrumento de 2004 a 2025 (A) e distribuição percentual por instrumento em 2022 (B) e 2025 (C).	11
2.2	Grafico de velas	12
2.3	Índice de força relativa no GBP/USD	19
2.4	Bandas de Bollinger e %B no EUR/USD.	21
2.5	Fluxo de dinheiro de Chaikin do WTI.	22
2.6	Intervalo verdadeiro médio do S&P500	24
2.7	Oscilador de momento de Chande no EUR/USD	25
2.8	Oscilador Estocástico no EUR/USD.	27
2.9	Índice de Movimento Direccional no GBP/USD.	29
2.10	Comparação entre volatilidades no par USD/JPY	40
2.11	Ilustração do VaR de $\alpha = 5\%$ de uma distribuição normal de retornos.	43
2.12	Ilustração de drawdowns no valor de um portfólio ao longo de um período de tempo	46
2.13	Diagrama de uma rede neural com uma camada oculta	54
2.14	Esquema do processo de treinamento de uma rede neural artificial, ilustrando o fluxo feed-forward, o cálculo do erro e o ajuste iterativo dos pesos por retropropagação.	56
2.15	Exemplo de particionamento recursivo, árvore correspondente e superfície de predição.	58
2.16	Curva ROC	64
2.17	(a) Grid search; (b) Random Search.	68
2.18	Validação cruzada k-fold clássica vs validação cruzada de série temporal.	69
2.19	Validação Cruzada com: (a) Janela Expandida; (b) Janela Deslizante.	70
4.1	Preços de fechamento e retornos aritméticos para os pares.	89
4.2	Histograma dos retonos aritméticos dos pares de moeda.	90
4.3	Heat Map da correlação entre os retornos diários dos activos.	91
4.4	Relação entre retornos e retornos do alvo sob diferentes condições do índice de força relativa e oscilador estocástico do USD/JPY.	92
4.5	Relação entre retornos e retornos do alvo sob diferentes condições do índice de força relativa e oscilador estocástico do EUR/USD.	92
4.6	Relação entre retornos e retornos do alvo sob diferentes condições do índice de força relativa e oscilador estocástico do GBP/USD.	93
4.7	Retornos médios por activo e periodo.	94
4.8	EUR/USD: (a) Correlação entre as variáveis independentes seleccionadas; (b) Correlação entre as variáveis independentes seleccionadas e o alvo	95
4.9	GBP/USD: (a) Correlação entre as variáveis independentes seleccionadas; (b) Correlação entre as variáveis independentes e o alvo	95

4.10	Correlação entre as variáveis independentes seleccionadas para os modelos do USD/JPY	96
4.11	Correlações positivas e negativas entre as variáveis seleccionadas para os modelos do USD/JPY x Alvo	96
4.12	Matrizes de confusão dos modelos do GBP/USD: (a) Regressão logística penalizada via L1; (b) Árvore de decisão; (c) Rede neural artificial.	99
4.13	Matrizes de confusão dos modelos do EUR/USD: (a) Regressão logística penalizada via L1; (b) Árvore de decisão; (c) Rede neural artificial.	99
4.14	Matrizes de confusão dos modelos do USD/JPY: (a) Regressão logística penalizada via L1; (b) Árvore de decisão; (c) Rede neural artificial.	99
4.15	Curvas ROC dos modelos do GBP/USD.	100
4.16	Curvas ROC dos modelos do EUR/USD.	100
4.17	Curvas ROC dos modelos do USD/JPY.	100
4.18	Desempenho da validação cruzada com janela expansiva no algoritmo RL-L para o par EUR/USD.	105
4.19	Desempenho da validação cruzada com janela expansiva no algoritmo AD-C para o par EUR/USD.	106
4.20	Desempenho da validação cruzada com janela expansiva no algoritmo RNA - BP para o par EUR/USD.	106
4.21	Desempenho da validação cruzada com janela expansiva no algoritmo RL-L para o par GBP/USD.	107
4.22	Desempenho da validação cruzada com janela expansiva no algoritmo AD-C para o par GBP/USD.	108
4.23	Desempenho da validação cruzada com janela expansiva no algoritmo RNA - BP para o par GBP/USD.	108
4.24	Desempenho da validação cruzada com janela expansiva no algoritmo RL-L para o par USD/JPY.	109
4.25	Desempenho da validação cruzada com janela expansiva no algoritmo AD-C para o par USD/JPY.	110
4.26	Desempenho da validação cruzada com janela expansiva no algoritmo RNA - BP para o par USD/JPY.	110
4.27	Comparação da curva do retorno diário acumulado das negociações com e sem a aplicação da estratégia baseada nas previsões da rede neural para o par GBP/USD.	112
4.28	Desempenho da Estratégia no GBP/USD: (a) Retorno por ano; (b) Frequência absoluta dos sinais; (c) Drawdowns; (d) Distribuição dos retornos.	113
4.29	Comparação do retorno diário acumulado das negociações com e sem a aplicação da estratégia baseada nas previsões da regressão logística L1 para o par USD/JPY.	115
4.30	Desempenho da estratégia híbrida e/ou modelo puro no USD/JPY: (a) Retorno por ano; (b) Frequência de sinais; (c) Drawdowns; (d) Distribuição dos retornos.	116
4.31	Comparação do retorno diário acumulado das negociações com e sem a aplicação da estratégia baseada nas previsões da rede neural para o par EUR/USD. . .	117
4.32	Desempenho da Estratégia no EUR/USD: (a) Retorno por ano; (b) Frequência absoluta dos sinais.	118
4.33	Desempenho da Estratégia no EUR/USD: (a) Drawdown; (b) Distribuição dos retornos.	118

4.34	Evolução temporal do coeficiente de HURST por taxa de variação do USD/JPY.	138
4.35	Evolução temporal do coeficiente de HURST por taxa de variação do GBP/USD.	138
4.36	Evolução temporal do coeficiente de HURST por taxa de variação do EUR/USD.	139
4.37	Evolução temporal da probabilidade estacionariedade do USD/JPY por taxa de variação.	139
4.38	Evolução temporal da probabilidade estacionariedade do EUR/USD por taxa de variação.	140
4.39	Evolução temporal da probabilidade estacionariedade do GBP/USD por taxa de variação.	140
4.40	Desempenho da validação cruzada para a regressão logística ($L1$) no par EUR/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra. . . .	141
4.41	Desempenho da validação cruzada para a árvore de decisão no par EUR/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.	141
4.42	Desempenho da validação cruzada para a rede neural no par EUR/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.	141
4.43	Desempenho da validação cruzada para a regressão logística ($L1$) no par GBP/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra. . . .	142
4.44	Desempenho da validação cruzada para a árvore de decisão ($L1$) no par GBP/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra. . . .	142
4.45	Desempenho da validação cruzada para a rede neural no par GBP/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.	142
4.46	Desempenho da validação cruzada para a regressão logística ($L1$) no par USD/JPY: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra. . .	143
4.47	Desempenho da validação cruzada para a árvore de decisão ($L1$) no par USD/JPY: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra. . .	143
4.48	Desempenho da validação cruzada para a rede neural no par USD/JPY: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.	143
4.49	Distribuição das acurácias no bootstrap por blocos moveis dos algoritmos do par EUR/USD.	144
4.50	Distribuição das acurácias no bootstrap por blocos moveis dos algoritmos do par GBP/USD.	144
4.51	Distribuição das acurácias no bootstrap por blocos moveis dos algoritmos do par USD/JPY.	144
4.52	Heat Map da correlação entre as variáveis do par EUR/USD.	145
4.53	Heat Map da correlação entre as variáveis do par GBP/USD.	145
4.54	Heat Map da correlação entre as variáveis do par USD/JPY.	145
4.55	Árvore de decisão do par EUR/USD.	146
4.56	Árvore de decisão do par GBP/USD.	146
4.57	Árvore de decisão do par USD/JPY.	146

4.58	Curvas dos retornos acumulados dos EUR/USD.	147
4.59	Curvas dos retornos acumulados dos GBP/USD.	147

Lista de Tabelas

2.1	Classificação do coeficiente de variação	32
2.2	Matriz de confusão em classificação binária.	62
2.3	Tabela de contingência para o Teste de McNemar	66
3.1	Descrição das variáveis	77
3.2	Base de treino,	78
3.3	Base de teste.	78
3.4	Hiperparâmetros para optimização do aprendizado dos algoritmos	83
3.5	Crítérios de selecção de modelos	86
3.6	Procedimento de decisão para selecção de modelos	86
4.1	Estatísticas Descritivas	89
4.2	Desempenho geral dos algoritmos na previsão da direcção dos retornos diários.	97
4.3	Testes de McNemar pareado entre os algoritmos.	101
4.4	Testes de Delong pareado entre os algoritmos.	101
4.5	Testes de McNemar pareado para os algoritmos.	102
4.6	Testes de Delong pareado para os algoritmos.	102
4.7	Testes de McNemar pareado para os algoritmos.	102
4.8	Testes de Delong pareado para os algoritmos.	103
4.9	Avaliação do desempenho dos modelos do EURUSD utilizando bootstrap por blocos moveis.	103
4.10	Avaliação do desempenho dos modelos do GBPUSD utilizando bootstrap por blocos moveis (%).	104
4.11	Avaliação do desempenho dos modelos do USDJPY utilizando bootstrap por blocos moveis (%)	104
4.12	Medidas de desempenho da rede neural e da Estratégia no GBP/USD.	114
4.13	Medidas de desempenho da regressão logística L1 e da estratégia no par USD/JPY.	117
4.14	Medidas de desempenho da rede neural e da estratégia no par EUR/USD.	119
4.15	Coeficientes da regressão logística penalizada via L1 optimizada para USDJPY, EURUSD e GBPUSD.	148
4.16	Parâmetros Utilizados nas Estratégias de Negociação.	148

Capítulo 1

INTRODUÇÃO

Foreign Exchange ou mercado de câmbio, também conhecido como Forex, é o maior mercado de câmbio de moedas estrangeiras do mundo, com 9,6 trilhões de dólares em volume medio diário de negociação, e a negociação de moedas decorre 24 horas por dia, cinco dias por semana (Archer, 2010; Bank for International Settlements, 2025). De acordo com Levinson (2014), o mercado de Forex é muito complexo e volátil, sendo frequentemente comparado a uma caixa preta, devido à sua natureza desconhecida e às grandes flutuações nas taxas de câmbio.

Ozturk et al. (2016) explicam que, os períodos de negociação são divididos em quatro grandes zonas horárias, a australiana, a asiática, a europeia e a norte-americana. Cada uma destas zonas apresenta horários específicos de abertura e fechamento.

Neste mercado, o dólar norte-americano (USD) é a moeda de referência para a maioria das negociações realizadas através de pares de moedas. Segundo Anastasakis (2009), o mercado é dividido em majors, cross-rates e exotics. Os majors, como indica Huang (1999), são os pares mais negociados, sempre cotados em relação ao USD. Os cross-rates envolvem moedas diferentes do USD, enquanto os exotics englobam moedas menos comuns, mas ainda cotadas em relação ao dólar (Masrya et al., 2013).

Cada par de moedas tem seu próprio preço de abertura, preço mais alto, preço mais baixo e preço de fechamento com base na sessão de negociação (trading). Adariani (2012) explica que, por razões de segurança, não é possível que uma pessoa acesse directamente o mercado Forex para comprar moedas. É necessário utilizar um intermediário, ou corrector, que tenha acesso ao mercado e possa realizar a compra directamente.

No mercado Forex, o participante tem duas opções: comprar ou vender moedas, buscando lucrar com a diferença entre as taxas de câmbio de entrada e saída da operação. Como no mercado de acções, o mercado Forex também depende de eventos que afectam a economia de um país. Por isso, o operador de mercado precisa entender o comportamento do mercado antes de realizar suas negociações.

Autores como Malkiel (2003), com a hipótese da caminhada aleatória, e Jensen (1978), com a hipótese do mercado eficiente, defendiam por muito tempo que prever a direcção dos preços no mercado financeiro era uma tarefa praticamente impossível. No entanto, nos últimos anos, pesquisadores como Cuomo et al. (2018) apresentaram várias ideias únicas usando aprendizado de máquina (*machine learning*) para realizar essa tarefa. De modelos estatísticos a modelos de aprendizado profundo (*deep learning*), diferentes modelos, juntamente com modelos híbridos, surgiram nessa ocasião (Yu, Wang & Lai, 2005; Ince & Trafalis, 2006; Hajizadeh et al., 2019). Esses modelos não são apenas diversos, mas também diferentes entre si.

Dado o contexto apresentado, o presente trabalho tem como objectivo investigar a aplicação de técnicas de machine learning na optimização de retornos financeiros no mercado Forex.

1.1 Problematização

Os investidores no mercado de Forex, desde institucionais até individuais, buscam constantemente metodologias eficazes para optimizar os seus retornos financeiros através da gestão e mitigação de risco, num ambiente de mercado dinâmico.

Métodos tradicionais de análise, como a análise técnica e a análise fundamental, tem suas limitações na previsão da direcção dos preços, devido à complexidade e à natureza não linear dos dados. Menkhoff et al. (2018), destaca que a análise técnica, se baseia em padrões históricos de preços e volume de negociação, o que pode falhar em capturar eventos inesperados e mudanças rápidas de mercado. Por outro lado, Fama e French (1996) abordam qua a análise fundamental, avalia o valor intrínseco dos activos com base em indicadores económicos e financeiros, pode ser menos eficaz em mercados altamente voláteis e dinâmicos como o Forex.

Vários estudos indicam que os modelos de machine learning podem aprimorar previsões no mercado financeiro (Rasekhschaffe & Jones, 2019; Gu et al., 2020). Pelo que, investidores profissionais adoptaram essas técnicas para negociação e gestão de portfólios (Barclaycobertura, 2018; Petropoulos et al., 2022).

É neste contexto que o presente estudo pretende investigar como os modelos de machine

learning podem ser aplicados para melhorar a previsão de retornos no mercado Forex e assim otimizar os resultados financeiros dos investidores. Pelo que propõe-se a responder a seguinte pergunta:

Até que ponto os modelos de machine learning podem aprimorar a previsão, visando a maximização de retornos financeiros no mercado de Forex?

1.2 Objectivos

1.2.1 Geral

Otimizar os retornos financeiros de investimentos realizados no mercado de Forex, através de uso de técnicas de machine learning.

1.2.2 Específicos

- Realizar uma análise exploratoria dos dados, para cada par de moedas seleccionado;
- Aplicar os modelos de regressão logística penalizada via L1, árvore de decisão e redes neurais para prever a direcção dos retornos de investimentos no mercado Forex;
- Seleccionar dentre os três modelos aplicados, aquele que melhor se ajusta aos dados históricos dos pares de moedas considerados;
- Desenvolver e testar uma estratégia de negociação que utilize as previsões do modelos, incorporando critérios de gestão de risco, para maximizar retornos no mercado Forex.

1.3 Justificação

A incorporação de técnicas de machine learning na análise de dados financeiros vem se destacando como uma alternativa promissora para superar as limitações das abordagens convencionais. Modelos de machine learning tem demonstrado capacidade de identificar padrões ocultos e prever retornos de forma mais eficaz (Fischer, 2018). As redes neurais artificiais, por exemplo, tem demonstrado desempenho superior em várias tarefas de previsão financeira, ao detectarem padrões complexos e não lineares, esses métodos aproveitam a força do algoritmo (Masini et al., 2020).

Este estudo é relevante por explorar de que forma diferentes abordagens de aprendizado de máquina (machine learning) se comportam quando aplicadas ao mercado de Forex, com-

parando modelos como a regressão logística penalizada via L1, a árvore de decisão e as redes neurais. Com o objectivo de maximizar os retornos gerados por estes modelos, serão aplicadas de técnicas de gestão de risco, como o VaR, entre outras técnicas. Ao validar estas metodologias com dados históricos de pares de moedas, o trabalho procura oferecer contributos teóricos e práticos que possam ser utilizados tanto por profissionais do mercado financeiro como por investidores individuais, com o objectivo de melhorar as suas estratégias de investimento e análise financeira.

Assim, ao investigar a eficiência de modelos de machine learning no contexto do mercado de Forex, prevê-se que o estudo contribuirá para o desenvolvimento de estratégias mais precisas de previsão da direcção dos retornos, possibilitando uma tomada de decisão mais informada para investidores neste mercado marcado por alta volatilidade e incerteza.

1.4 Estrutura do trabalho

O presente trabalho encontra-se estruturado em cinco capítulos, sendo os subsequentes capítulos destacados a seguir.

O Capítulo 2 (**Revisão da Literatura**) apresenta o enquadramento teórico do estudo, reunindo os principais conceitos e fundamentos teóricos que sustentam as técnicas aplicadas com o propósito de atingir o objectivo geral do estudo.

O Capítulo 3 (**Material e Métodos**) descreve o material usado para a realização do trabalho e os métodos usados para obtenção dos resultados.

No Capítulo 4 (**Resultados e Discussão**) são apresentados os resultados obtidos, procedendo-se à sua interpretação e discussão à luz da literatura e dos objectivos propostos.

Por fim, o Capítulo 5 (**Conclusões e Recomendações**) reúne as principais conclusões do estudo e recomendações para investigações futuras. Adicionalmente elementos complementares, tais como as referências bibliográficas e apêndice.

Capítulo 2

REVISÃO DA LITERATURA

2.1 Conceitos básicos

Os mercados financeiros constituem estruturas essenciais para o funcionamento da economia, permitindo a transferência eficiente de recursos e a formação de preços de activos. Segundo Gitman e Joehnk (2005), são conjuntos de instituições (como bancos, correctoras, fundos de investimento e bolsas de valores) e instrumentos (como acções, índices, moedas e produtos primários ou *commodities*) que proporcionam diferentes formas de utilização e alocação de recursos financeiros.

Oliveira (1983) acrescenta que, além de viabilizar a transferência de recursos dos fornecedores para os tomadores finais, os mercados financeiros garantem a liquidez dos activos, permitindo que sejam rapidamente convertidos em dinheiro ou outros instrumentos financeiros, assegurando assim a eficiência e dinamismo das transações.

Outro conceito importante é o de investimento. De acordo com Blanchard (2017), o investimento contribui para o aumento dos recursos produtivos de uma economia, ampliando sua capacidade de produção. Menezes (2011) ressalta que o investimento busca maximizar riqueza e retornos futuros, geralmente em horizonte de longo prazo, enquanto Gitman e Joehnk (2005) definem o investimento como a disponibilização antecipada de recursos para aumentar ou preservar seu valor.

Graham (2016) estabelece uma distinção crucial entre investimento e especulação. Para o autor, uma operação de investimento é aquela que, após análise profunda, promete segurança do capital principal e um retorno adequado, pelo que operações que não atendem a esses critérios são consideradas especulativas.

2.2 Mercado de Forex

2.2.1 Breve histórico

Embora a história do mercado de Forex (ou mercado de câmbio) seja quase tão longa quanto a própria civilização, basta recuar até 1944 para entender as bases da estrutura moderna desse mercado.

A consolidação moderna do mercado de Forex ocorreu apenas ao final da segunda guerra mundial, quando o acordo de Bretton Woods, firmado em 1944 por 44 países, estabeleceu um sistema monetário internacional baseado na fixação das taxas de câmbio (Kritzer, 2012). Nesse arranjo, o dólar norte-americano assumiu o papel de moeda principal, com paridade directa ao ouro, enquanto as demais moedas eram fixadas ao dólar. O objectivo central era reconstruir a ordem económica global, reduzir assimetrias cambiais e criar um ambiente internacional de estabilidade monetária capaz de fomentar o comércio e a recuperação das economias no pós-guerra.

A estrutura de Bretton Woods foi fundamental para a expansão do comércio internacional nas décadas seguintes, entretanto, suas limitações tornaram-se evidentes à medida que os fluxos financeiros globais aumentaram e a convertibilidade rígida do dólar passou a pressionar as reservas norte americanas. A crescente discrepância entre o volume de dólares circulando internacionalmente e a quantidade de ouro disponível para garantir o dólar gerou tensões estruturais que culminaram na suspensão da conversibilidade do dólar em ouro em 1971 (Kritzer, 2012).

O sistema foi oficialmente encerrado em 1973, dando origem ao regime de câmbio flutuante, frequentemente denominado Bretton Woods II, no qual o valor das moedas passou a ser determinado pelas forças de oferta e demanda no mercado (Kritzer, 2012). Esse marco redefiniu o funcionamento do mercado cambial, tornando-o mais dinâmico, descentralizado e sensível aos fluxos globais de capitais. Desde então, o Forex expandiu-se rapidamente devido à liberalização financeira e ao avanço das tecnologias de negociação eletrónica, consolidando-se como o maior e mais líquido mercado financeiro do mundo, com volume diário superior a trilhões de dólares (Kritzer, 2012).

O acordo de Bretton Woods também resultou na criação de instituições fundamentais para supervisionar o novo sistema financeiro e garantir a liquidez na economia global, como o Fundo Monetário Internacional (FMI) e o Banco Internacional para Reconstrução e Desenvolvimento (BIRD), actualmente conhecido como Banco Mundial (Kritzer, 2012).

2.2.2 Principais participantes

O ambiente cambial global é formado por uma ampla rede de participantes com diferentes motivações e responsabilidades, geralmente referidos como negociantes (*traders*). Entre os agentes principais estão corporações multinacionais, bancos comerciais e centrais, instituições financeiras, fundos de investimento, correctoras e investidores individuais (Kritzer, 2012; Archer, 2010). Cada grupo participa do mercado por razões distintas, enquanto investidores individuais muitas vezes buscam ganhos especulativos de curto prazo, empresas que actuam em comércio exterior utilizam o mercado para protecção contra riscos cambiais, gestão de exposição cambial e otimizando fluxos financeiros (Dolan, 2011; Ponsi, 2007). A diversidade de agentes e estratégias é uma das características que tornam o mercado de Forex altamente líquido e operacional 24 horas por dia, sendo também um mercado altamente competitivo e dinâmico, reflectindo constantemente factores macroeconómicos e geopolíticos (Bank for International Settlements, 2013; Babypips, 2015).

2.2.3 Mecanismos de negociação de moedas

O funcionamento do mercado cambial apresenta características únicas que o diferenciam de outros mercados financeiros. Primeiramente, não existe um centro físico de negociação, as operações são realizadas de forma descentralizada no ambiente de mercado de balcão (OTC), isto é, fora da bolsa de valores tradicionais, permitindo que participantes de diferentes regiões interajam directamente entre si a qualquer hora do dia (Kritzer, 2012; Dolan, 2011). Essa estrutura descentralizada contribui para a alta liquidez do mercado e para a possibilidade de negociações contínuas, 24 horas por dia, cinco dias por semana (Bank for International Settlements, 2025).

Todas as negociações envolvem pares de moedas, ou seja, ao comprar uma moeda, o investidor vende simultaneamente a moeda de referência, conhecida como moeda de cotação. Essa característica implica que a valorização ou desvalorização de uma moeda só pode ser avaliada em relação a outra, tornando a análise do mercado relativa e não absoluta (Kritzer, 2012).

Um aspecto central no mercado Forex é a alavancagem, que permite ao trader operar com valores muito superiores ao capital disponível, multiplicando tanto o potencial de lucro quanto o risco de perda. Por exemplo, uma alavancagem de 1:100 permite controlar uma posição de \$100,000 com apenas \$1,000 de capital próprio. Esse mecanismo funciona em conjunto com a margem, que é o valor mínimo necessário para manter a posição aberta. A gestão correcta da margem e da alavancagem é fundamental, pois movimentos desfavoráveis no mercado podem resultar em chamadas de margem ou liquidação automática de posições (Kritzer, 2012; Hull, 2018).

2.2.4 Moedas negociadas e principais pares

No mercado de Forex, as moedas são geralmente classificadas em principais (majors) e exóticas, de acordo com sua liquidez e relevância no comércio internacional. As moedas principais, como o dólar norte-americano (USD), euro (EUR) e iene japonês (JPY), dominam a maior parte das transações globais, sendo amplamente aceitas e utilizadas como referência para operações cambiais (BIS, 2025).

Os pares cambiais podem ser divididos em pares principais, que envolvem o dólar norte-americano, e pares cruzados, nos quais o dólar não está presente. Essa classificação auxilia na compreensão da dinâmica de liquidez e na escolha de estratégias de negociação (Kritzer, 2012). Entre os pares mais negociados globalmente estão EUR/USD, USD/JPY e GBP/USD, apresentando elevada liquidez, estabilidade relativa e ampla disponibilidade de informações (BIS, 2025; Dolan, 2011).

Além das majors, algumas moedas exóticas e de mercados emergentes vêm ganhando relevância, oferecendo oportunidades de diversificação, embora apresentem maior volatilidade e spreads mais amplos. Dessa forma, o mercado de Forex combina pares altamente líquidos e populares com alternativas menos convencionais, permitindo diferentes estratégias de negociação e gestão de risco (BIS, 2025).

2.2.5 Processos de negociação e principais tipos de ordem

O acesso ao mercado de Forex por investidores individuais é feito essencialmente por meio de correctoras que fornecem plataformas eletrônicas de negociação. Essas plataformas permitem visualizar cotações em tempo real, gráficos interactivos, indicadores técnicos, histórico de operações, ferramentas de análise de risco, além de permitirem o desenvolvimento de algoritmos e a automatização das negociações, facilitando a tomada de decisões (Kritzer, 2012; Dolan, 2011). Já Investidores institucionais, como grandes bancos, fundos de investimento e corporações multinacionais tem acesso directo ao mercado interbancário.

Cada par de moedas apresenta dois preços simultâneos: o preço de compra (*bid*), correspondente ao valor pelo qual o participante no mercado pode vender a moeda base, e o preço de venda (*ask*), correspondente ao valor pelo qual a pode adquirir. A diferença entre esses preços é conhecida como *spread*, sendo um dos principais custos de transação. A menor variação possível no preço de um par é chamada de percentual em pontos (*pip*) (Ponsi, 2007; Kritzer, 2012).

Para iniciar uma operação, o participante do mercado deve inserir uma ordem. As principais categorias incluem:

- **Ordem de mercado:** executada imediatamente ao preço vigente, garantindo rapidez na entrada ou saída da posição.
- **Ordem de realizar lucro:** instrui o fechamento automático da posição quando o preço atinge um nível predeterminado de lucro, permitindo ao trader concretizar ganhos sem monitoramento constante.
- **Ordem de parar a perda:** estabelece um limite de perda, encerrando a posição automaticamente quando atingido, sendo crucial para protecção do capital investido (Dolan, 2011; Ponsi, 2007).

As posições podem ser abertas de duas formas, compra da moeda base e venda da moeda cotada (*long*) ou venda da moeda base e compra da moeda cotada (*short*). A primeira posição é lucrativa quando a moeda base se valoriza em relação à moeda cotada, enquanto a segunda posição gera lucro quando a moeda base se desvaloriza. O fechamento da posição ocorre por meio da operação inversa, realizando lucro ou prejuízo conforme a variação do preço do par de moeda (Kritzer, 2012; Dolan, 2011).

Além das ordens básicas, os operadores de mercado financeiro podem utilizar ordens condicionais, ordens de limite, ordens de parada de limite (Kritzer, 2012; Murphy, 2015).

2.2.6 Instrumentos de negociação de moedas

Os instrumentos de negociação de moedas constituem a base operacional do mercado de Forex, permitindo que agentes económicos transfram, protejam ou ampliem sua exposição às flutuações das taxas de câmbio. A literatura de finanças internacionais destaca que tais instrumentos são essenciais para a eficiência do mercado de Forex, pois viabilizam tanto operações comerciais quanto estratégias de investimento e cobertura neste ambiente global, descentralizado e de elevada liquidez (Mishkin, 2015; Madura, 2020). Além de possibilitarem a conversão directa entre moedas, esses instrumentos desempenham papel relevante na formação de expectativas, na descoberta de preços e na gestão de riscos cambiais.

A literatura especializada agrupa os instrumentos cambiais em operações à vista e derivativos cambiais, sendo este segundo representado por contratos a termo, contratos futuros, contratos por diferença, opções cambiais e *swaps*, cujos valores derivam do par de moedas. Cada modalidade apresenta características distintas quanto ao risco de contraparte, custos operacionais, estrutura de liquidação e finalidade económica (Kritzer, 2011; Dolan, 2011).

Assim, investidores individuais, empresas, instituições financeiras, gestores de fundos e organismos reguladores utilizam diferentes combinações desses instrumentos conforme suas

necessidades de cobertura, especulação ou arbitragem, integrando-os às suas estratégias de gestão de risco e de alocação de capital. São estes:

- **Operações à vista (*Spot*):** Referem-se à compra ou venda imediata de uma moeda segundo sua cotação actual, com liquidação geralmente em dois dias úteis. Esse tipo de contrato representa o mecanismo padrão de negociação para negociadores individuais e institucionais, servindo como referência para a formação de preços de derivativos cambiais (Kritzler, 2011; Dolan, 2011). O mercado spot possui alta liquidez e é influenciado principalmente por fluxos comerciais, notícias macroeconômicas e intervenções monetárias.
- **Contratos a termo (*Forward*):** São instrumentos financeiros derivativos, negociados em mercado de balcão (OTC) e especificam a taxa de câmbio e a data futura de liquidação. Por serem personalizados, permitem adequação exata às necessidades de empresas e instituições financeiras que buscam protecção contra a volatilidade cambial. Esses contratos não exigem pagamento inicial e apresentam risco de contraparte, já que não são negociados em bolsa (Kritzler, 2011; Archer, 2010).
- **Contratos futuros (*Futures*):** São derivativos padronizados, negociados em bolsas, e liquidados por meio de câmaras de compensação, o que reduz substancialmente o risco de contraparte. Diferem dos contratos a termo pela padronização dos lotes e datas de vencimento, além da obrigatoriedade de margens e ajustes diários. São amplamente utilizados para estratégias de arbitragem, especulação alavancada e protecção cambial estruturada (Kritzler, 2011; Ponsi, 2007).
- **Contratos por diferença (*CFDs*):** São instrumentos derivativos de balcão que permitem ao investidor especular sobre a variação do preço de um activo subjacente sem necessidade de posse efectiva desse activo. No mercado cambial, os CFDs replicam movimentos de pares de moedas à vista, possibilitando a abertura de posições longas ou curtas com elevada alavancagem e baixo requisito de margem, características que os tornam populares entre traders de curto prazo (Norman, 2009; Cartridge & Jessen, 2011). Os CFDs são instrumentos oferecidos e providos por correctoras e plataformas de negociação que actuam como contraparte directa do cliente.
- **Opções cambiais (*Currency options*):** São derivativos que concedem ao comprador o direito, mas não a obrigação, de comprar (*call*) ou vender (*put*) uma moeda por um preço predeterminado (*strike*). São instrumentos flexíveis, adequados para hedge (cobertura) assimétrica, estratégias avançadas de protecção e operações de volatilidade (Dolan, 2011; Kritzler, 2011).

- **FX swaps:** Estes derivativos, consistem na troca simultânea de duas moedas no mercado a vista, seguida da recompra ou revenda das mesmas moedas por meio de um contrato a termo com data futura acordada. Os FX Swaps se limitam à combinação de uma operação á vista e outra a termo, geralmente com prazos curtos. São amplamente utilizados por bancos centrais, instituições financeiras e grandes corporações para gestão de liquidez, ajuste de posições cambiais e protecção contra oscilações de curto prazo na taxa de câmbio (Kritzer, 2011; Dolan, 2011).

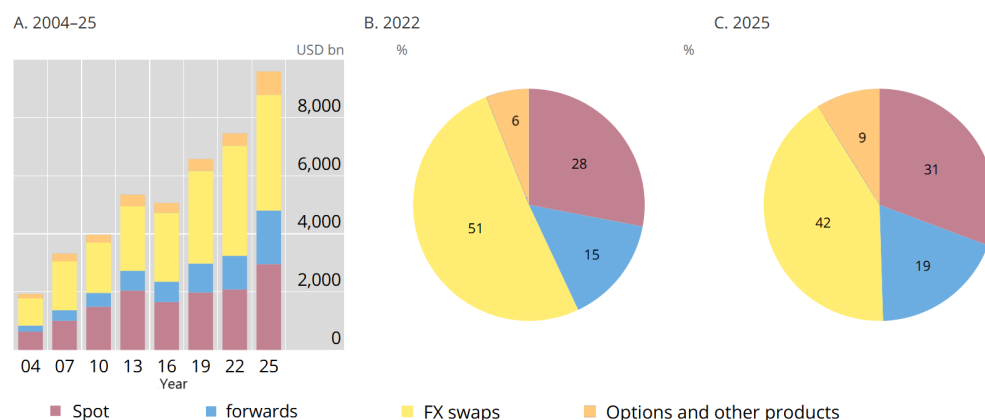


Figura 2.1: Volume médio diário negociado do mercado Forex por instrumento de 2004 a 2025 (A) e distribuição percentual por instrumento em 2022 (B) e 2025 (C).

Fonte: Adaptado de BIS (2025).

2.2.7 Análise fundamental e análise técnica

A capacidade de antecipar movimentos de preços no mercado de Forex é um dos pilares da actividade de negociação (*trading*) e gestão de risco financeiro. Neste mercado, os participantes recorrem a diferentes abordagens analíticas para fundamentar suas decisões. A literatura especializada distingue principalmente dois grandes métodos, análise fundamental e análise técnica (Dolan, 2011; Kritzer, 2012; Murphy, 1999).

- **Análise fundamental:** Esta abordagem concentra-se nas causas subjacentes que influenciam o movimento dos preços das moedas, examinando factores macroeconómicos, políticas monetárias e fluxos internacionais de capital. Indicadores como produto interno bruto (PIB), inflação, níveis de taxas de juros, balança comercial e políticas fiscais são elementos centrais da análise fundamental (Dolan, 2011; Kritzer, 2012). A literatura destaca que eventos geopolíticos e decisões de bancos centrais, como ajustes nas taxas de referência, podem gerar efeitos significativos sobre a cotação das moedas, pelo que a análise fundamental é uma das abordagens usadas para antecipar movimentos sustentados de preços (Archer, 2010; Chen et al., 2019).

- **Análise técnica:** Concentra-se na avaliação de padrões históricos de preços OHLC (Abertura, Máxima, Mínima, Fechamento) para antecipar movimentos futuros, baseando-se na premissa de que toda a informação relevante está incorporada no preço, de que os preços tendem a mover-se em tendências e de que padrões passados tendem a se repetir (Murphy, 1999; Kritzer, 2012). Os principais instrumentos desse tipo de análise incluem o estudo de padrões gráficos, sendo o mais difundido o gráfico de velas (candles ticks). Adicionalmente, utilizam-se indicadores técnicos e o reconhecimento de formações gráficas (Dolan, 2011; Stock & Watson, 2015).

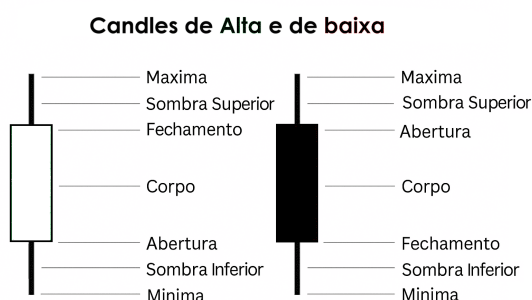


Figura 2.2: Gráfico de velas

2.2.8 Relações intermercados

As relações intermercados constituem uma das principais referenciais para compreender como diferentes segmentos do sistema financeiro global se influenciam mutuamente. O mercado de Forex, por sua dimensão e sensibilidade a movimentos macroeconómicos, encontra-se frequentemente associado ao comportamento de outros mercados, como por exemplo, índices e produtos primários (commodities), que reflectem expectativas agregadas sobre actividade económica, risco e liquidez internacionais (Murphy, 1991; Fabozzi et al., 2014).

Do ponto de vista macroeconómico e fundamentalista, as oscilações cambiais respondem a factores como crescimento económico, política monetária, inflação e aversão global ao risco. Esses factores influenciam simultaneamente diferentes segmentos dos mercados financeiros, dando origem a interdependências entre mercados de acções, mercados cambiais e outros activos financeiros, amplamente documentadas na literatura (Mishkin, 2016; Obstfeld & Rogoff, 1996).

No contexto europeu, estudos que analisam índices de acções regionais mostram que o Euro Stoxx 50, enquanto indicador agregado do desempenho das principais economias da zona do euro, contém informação relevante sobre o sentimento económico e financeiro da região. Variações nesse índice tem sido associadas a movimentos em pares cambiais relacionados ao euro, reflectindo mudanças nas expectativas de crescimento, estabilidade financeira e fluxos de capitais internacionais (Caporale, Hunter & Menla Ali, 2014).

Garcia e Rodrigues (2019) identificaram em seus estudos uma relação unidireccional entre o índice FTSE 100 e a taxa de câmbio USD/EUR. Ainda no mesmo estudo, foi detectada a existência de uma relação unidireccional entre a taxa de câmbio USD/GBP e os índices Euro STOXX 50 e FTSE 100.

De forma semelhante, o mercado de acções norte-americano, representado pelo índice S&P 500, desempenha um papel central no sistema financeiro global. A literatura aponta que o S&P 500 funciona como um barómetro da aversão global ao risco, influenciando os fluxos internacionais de capitais e, consequentemente, as taxas de câmbio. Em períodos de risk-on (maior apetite ao risco), a valorização do mercado de acções norte-americano tende a estar associada à depreciação do dólar face a moedas de economias avançadas, enquanto episódios de risk-off (maior aversão ao risco) conduzem frequentemente à valorização do dólar como activo de refúgio (Engel, Mark & West, 2015).

Fatum e Yamamoto (2016), em seus estudos mostram que choques negativos nos mercados de acções globais, incluindo quedas no S&P 500 ou no Nikkei 225, estão frequentemente associados à apreciação do iene, reflectindo realocações defensivas de capital em períodos de incerteza. Tsuji (2015) em seus estudos identificou movimentações estatisticamente significativas entre o índice Nikkei 225 e mudanças na taxa de câmbio dólar-iene, indicando uma interdependência dinâmica entre estes mercados.

O índice do dólar norte-americano (DXY) é amplamente utilizado na literatura como uma medida agregada da força do dólar face a um conjunto de moedas relevantes. Variações no DXY reflectem alterações nas condições monetárias dos Estados Unidos, diferenciais de juros e perceções globais de risco, exercendo influência directa sobre pares como EUR/USD e GBP/USD. Nesse sentido, estudos indicam que o DXY possui elevado poder explicativo para a dinâmica de curto e médio prazo das principais taxas de câmbio, sobretudo em ambientes de elevada volatilidade financeira (Verdelhan, 2018).

Os produtos primários (*commodities*) também ocupam um papel central na literatura de relações intermercado, uma vez que os seus preços afectam directamente as balanças comerciais, os termos de troca e as expectativas inflacionárias, influenciando simultaneamente mercados cambiais e de acções. Em particular, o preço do petróleo bruto, frequentemente representado pelo West Texas Intermediate (WTI), tem sido amplamente utilizado como indicador de choques globais de oferta e procura e das condições macroeconómicas internacionais. Estudos clássicos demonstram que variações no preço do petróleo estão associadas a movimentos sistemáticos nas taxas de câmbio, reflectindo ajustamentos nos fluxos comerciais e financeiros, bem como alterações nas expectativas de inflação e de política monetária (Chen & Rogoff, 2003).

2.3 Hipóteses do passeio aleatório e do mercado eficiente

A Hipótese do Passeio Aleatório e a Hipótese do Mercado Eficiente estão entre os conceitos mais debatidos na literatura financeira das últimas décadas. Apesar de amplamente conhecidas, ambas permanecem alvo de questionamentos e controvérsias quanto à sua validade prática (Fama, 1970; Fama & French, 1988; Lo & MacKinlay, 1988; Adnan et al., 2023).

A hipótese do passeio aleatório sugere que os preços dos activos financeiros seguem trajectórias aleatórias e imprevisíveis, de modo que não é possível extrair retornos excedentes a partir da observação de movimentos passados. Embora isso não exclua a possibilidade de ganhos por meio de informação privilegiada, implica que estratégias baseadas em séries históricas de preços não são capazes de gerar retornos anormais de forma consistente. Já a hipótese do mercado eficiente, de forma geral, sustenta que os mercados incorporam integralmente toda a informação disponível, e que os preços se ajustam de maneira imediata quando novas informações se tornam públicas. Contudo, na realidade do mercado, observa-se que alguns investidores reagem prontamente, enquanto outros preferem aguardar sinais mais consistentes, reagindo apenas quando uma tendência já está bem definida (Kondratenko & Kuperin, 2003).

No mercado cambial, pesquisas recentes reforçam essa perspectiva, indicando que a hipótese do mercado eficiente é um pilar da teoria financeira moderna, existem evidências de que os preços no Forex já reflectem rapidamente as informações disponíveis, o que limita a possibilidade de ganhos sistemáticos acima da média (Tiwari et al., 2024; Agyei et al., 2023).

Essas hipóteses possuem implicações relevantes para a teoria financeira e para a prática de investimentos. Por sua vez, desde os trabalhos de referência de Fama (1970), diversos estudos empíricos identificaram desvios em relação ao passeio aleatório. Muitos desses estudos encontraram evidências que desafiam a hipótese do passeio aleatório, sugerindo a existência de padrões previsíveis em determinados contextos (Fama & French, 1988; Lo & MacKinlay, 1988). Esses desvios indicam a presença de ineficiências no mercado financeiro, fenómeno em que os preços dos activos não reflectem completamente todas as informações disponíveis.

2.4 Indicadores técnicos

Os indicadores técnicos são ferramentas amplamente utilizadas na análise técnica de activos financeiros. Eles consistem em cálculos matemáticos aplicados a preços históricos, volume de negociação e outras variáveis de mercado, com o objectivo de identificar tendências, padrões

e possíveis pontos de reversão. Esses indicadores auxiliam analistas a tomar decisões fundamentadas, oferecendo sinais de compra ou venda com base em características estatísticas dos preços (Murphy, 1999).

Segundo Pring (2002), os indicadores técnicos não predizem o futuro, mas fornecem informações úteis sobre o comportamento do mercado. As subseções a seguir apresentam alguns dos indicadores técnicos.

Média móvel simples

As médias móveis são um tipo de método de suavização utilizado para reduzir ou eliminar a variação aleatória presente em dados coletados ao longo do tempo (Chatfield, 2003; Hyndman & Athanasopoulos, 2018). Quando aplicadas de forma adequada, essas técnicas revelam mais claramente a tendência subjacente, bem como padrões sazonais e cíclicos existentes nos dados.

Em aplicações financeiras, uma média móvel simples (ou simplesmente SMA, do inglês *Simple Moving Average*) de n períodos corresponde à média não ponderada dos n últimos pontos da série temporal (Chatfield, 2003). Seja a série temporal representada por y_t , a fórmula é dada por:

$$SMA_{n,t} = \frac{y_{t-(n-1)} + \cdots + y_{t-1} + y_t}{n} = \frac{1}{n} \sum_{i=0}^{n-1} y_{t-i}. \quad (2.1)$$

Como exemplo, a média móvel simples de quatro períodos é:

$$SMA_{3,t} = \frac{y_{t-3} + y_{t-2} + y_{t-1} + y_t}{4}. \quad (2.2)$$

A média móvel é, portanto, uma série temporal filtrada. Quando se calculam valores sucessivos, um novo dado entra na soma enquanto o mais antigo é descartado. Dessa forma, não é necessário recalcular a soma completa a cada instante, pois a média móvel simples pode ser obtida de forma recursiva (Hyndman & Athanasopoulos, 2018).

Médias móveis exponenciais

Os métodos de suavização exponencial são bastante populares para a construção de séries temporais suavizadas (Chatfield, 2003). Eles se assemelham às médias móveis ponderadas,

sendo, na verdade, um caso particular delas, em que observações mais recentes recebem pesos maiores do que observações mais antigas.

A característica distintiva da média móvel exponencial (ou *Exponential Moving Average* – EMA) é que os pesos decrescem de forma exponencial à medida que as observações ficam mais antigas. A seguir, apresenta-se o esquema de suavização exponencial simples (Hyndman & Athanasopoulos, 2018).

- **Suavização exponencial simples**

Na suavização exponencial simples, a EMA de n períodos para a série temporal y_t é calculada recursivamente por:

$$EMA_{n,t} = \alpha y_t + (1 - \alpha)EMA_{n,t-1}, \quad 3 \leq t \leq n. \quad (2.3)$$

Ao expandir esta relação, obtém-se a fórmula geral:

$$EMA_{n,t} = \alpha \sum_{j=0}^{t-1} (1 - \alpha)^j y_{t-j} + (1 - \alpha)^{t-1} S_{E2}, \quad 2 \leq t \leq n. \quad (2.4)$$

Aqui, $0 \leq \alpha \leq 1$ é o parâmetro de suavização, representando a intensidade de redução de peso. Um valor alto de α (próximo de 1) dá maior importância a variações recentes, reagindo rapidamente, mas descartando observações mais antigas mais depressa (Chatfield, 2003). Já valores baixos de α aumentam a suavização, atribuindo peso maior a valores passados e reduzindo a sensibilidade a mudanças recentes (Makridakis et al., 1998).

Outro aspecto é a inicialização da EMA. Diferentemente da SMA, não existe uma definição única para $EMA_{n,1}$, e portanto deve-se adotar um critério para definir o valor inicial. Uma prática comum é iniciar a EMA com a média simples dos primeiros n dados da série (Hyndman & Athanasopoulos, 2018). O impacto da escolha inicial tende a ser irrelevante quando t cresce, mas pode ser significativo nos primeiros períodos.

Além disso, a previsão com base na suavização exponencial simples pode ser expressa como:

$$EMA_{n,t+1} = \alpha y_t + (1 - \alpha)EMA_{n,t}, \quad (2.5)$$

ou ainda como:

$$EM A_{n,t+1} = EM A_{n,t} + \alpha e_t, \quad (2.6)$$

em que $e_t = y_t - EM A_{n,t}$ é o erro de previsão no período t .

Taxa de mudança

Este indicador inicia no conceito de momento (*momentum*), que é um indicador de taxa de mudança projectado para identificar a velocidade ou força de um movimento de preço (Murphy, 1999). Seu valor actual é simplesmente a diferença entre o preço actual e o preço de n períodos atrás:

$$\text{momentum}_t = p_t - p_{t-n} \quad . \quad (2.7)$$

Valores acima (abaixo) de zero indicam claramente que o preço (p_t) tem um momento de alta (baixa).

A taxa de mudança (ROC, do inglês *Rate of Change*) é um oscilador de momento que mede a variação percentual no preço de um activo em um determinado número de períodos, conforme representado pela fórmula:

$$\text{ROC}_t = \frac{p_t - p_{t-n}}{p_{t-n}} \times 100 \quad . \quad (2.8)$$

O valor do ROC no tempo t representa simplesmente o quanto o preço mudou em percentual desde n períodos atrás (Murphy, 1999). Sendo o ROC de $n = 1$, o calculo do retorno de um periodo.

Percentagem de amplitude de Williams

O Williams %R (*Williams percentage range*) ou percentagem de amplitude de Williams é um indicador de momento usado para medir níveis de sobrecompra e sobrevenda (Williams, 1979). Ele oscila no intervalo $[-100, 0]$, conforme a fórmula:

$$\text{WR}_{t,n} = -100 \cdot \frac{\max_{i \in [t-n+1, t]}(H_i) - C_t}{\max_{i \in [t-n+1, t]}(H_i) - \min_{i \in [t-n+1, t]}(L_i)} \quad . \quad (2.9)$$

Primeiro, como em muitos outros indicadores, deve-se escolher o número de períodos n a serem considerados. Então, $\max_{i \in [t-n+1, t]}(H_i)$ denota o maior valor de preço máximo (o mais alto) observado nos últimos n períodos. De forma semelhante, $\min_{i \in [t-n+1, t]}(L_i)$ representa o menor valor de preço mínimo no mesmo intervalo. Por sua vez, C_t indica o preço de fechamento no período t . Em termos de interpretação, as leituras do Williams %R entre aproximadamente -80% e -100% indicam que o activo está sobrevendido e leituras próximas a 0% a -20% sugerem que o activo está sobrecomprado.

Índice de força relativa

O Índice de Força Relativa (IFR) é um indicador de *momentum*, que mede a velocidade e magnitude dos movimentos de preços direcionais. Ele é usado para acompanhar a força e fraqueza histórica de um activo ou mercado com base nos preços de fechamento de um período de negociação recente (Wilder, 1978; Murphy, 1999).

O cálculo do IFR procede da seguinte forma. Primeiro, para cada período de negociação, uma mudança positiva U ou mudança negativa D é calculada a partir dos preços de fechamento p_t e p_{t-1} de acordo com:

$$U = \max(p_t - p_{t-1}, 0) = \begin{cases} p_t - p_{t-1}, & \text{se } p_t \geq p_{t-1}, \\ 0, & \text{caso contrário.} \end{cases} \quad (2.10)$$

$$D = \max(p_{t-1} - p_t, 0) = \begin{cases} p_{t-1} - p_t, & \text{se } p_{t-1} \geq p_t, \\ 0, & \text{caso contrário.} \end{cases} \quad (2.11)$$

As médias U e D são então calculadas usando uma média móvel exponencial de n períodos. A razão entre essas médias é o factor de força relativa:

$$RS = \frac{EMA_n(U)}{EMA_n(D)}. \quad (2.12)$$

Esse factor de força relativa é então convertido em um índice de força relativa que oscila entre 0 e 100:

$$IFR = 100 - \frac{100}{1 + RS}. \quad (2.13)$$

Inerente às médias móveis e métodos de suavização está o grau de liberdade na escolha do período de suavização n . Como o IFR utiliza a média móvel exponencial, essa liberdade de escolha depende do caso em questão. O período de suavização mais comum para usar o IFR é $n = 14$ (Murphy, 1999).

O índice de força relativa é amplamente empregado na análise técnica como uma ferramenta para identificar possíveis pontos de reversão na direcção dos preços. Entre as principais interpretações associadas a esse indicador, destaca-se sua aplicação na identificação de zonas de sobrecompra e sobrevenda (Murphy, 1999). Quando os preços de um activo apresentam movimentos abruptos de alta ou de baixa, o IFR tende a reflectir esses extremos, sinalizando que o activo pode estar temporariamente supervalorizado ou subvalorizado, respectivamente. A superação de determinados limiares por esse índice é geralmente interpretada como indicativo de que uma inversão de tendência pode estar iminente.

De acordo com Wilder (1978) e Murphy (1999), valores do IFR acima de 70 indicam sobrecompra, enquanto valores abaixo de 30 sugerem sobrevenda. A zona entre 30 e 70 é considerada neutra. Cruzamentos acima de 30 com inclinação positiva costumam sinalizar compra, enquanto cruzamentos abaixo de 70 com inclinação negativa indicam possível venda.

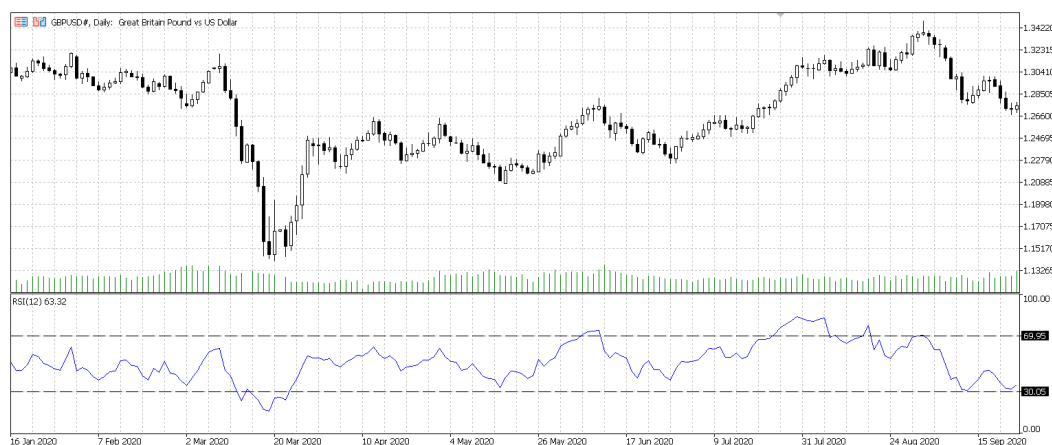


Figura 2.3: Índice de força relativa no GBP/USD

Fonte: Metatrader5

Percentual B das bandas de bollinger

O percentual B das bandas de bollinger é uma extensão das Bandas de Bollinger, primariamente é necessario definir o seu conceito. As bandas de Bollinger são utilizados para medir a volatilidade dos preços no mercado financeiro. Elas se baseiam no conceito estatístico de desvio padrão, o qual expressa o grau de dispersão dos preços em relação a uma média móvel (Bollinger, 2001).

Esse indicador é composto por três linhas principais:

- A **linha central**, que corresponde a uma média móvel simples dos preços;
- A **banda superior**, posicionada k desvios padrão acima da média dos preços;
- A **banda inferior**, posicionada k desvios padrão abaixo da média dos preços.

A fórmula geral das bandas é dada por:

$$B^{\pm} = MA_{t,n} \pm k\sigma \quad (2.14)$$

O desvio padrão é calculado conforme a seguinte expressão:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_{t-i+1} - MA_{t,n})^2}. \quad (2.15)$$

onde:

- p_{t-i+1} é o preço actual;
- $MA_{t,n}$ é a média móvel do preço em n períodos no tempo t ;
- σ representa o desvio padrão dos preços ao longo dos n períodos anteriores;
- k é um parâmetro que define o número de desvios padrão aplicados (frequentemente $k = 2$).

A principal funcionalidade das bandas de bollinger é ajustar-se automaticamente à volatilidade do mercado. Quando os preços se tornam mais voláteis, as bandas se expandem, e quando a volatilidade diminui, as bandas se contraem. Essa característica torna o indicador útil para identificar possíveis momentos de reversão de tendências. Além disso, quando ocorre uma contracção acentuada das bandas, isso pode anteceder movimentos mais expressivos no preço (Bollinger, 2001).

Introduzido por Blau (2001), o indicador percentual B das bandas de Bollinger ou simplesmente %B, é uma extensão que normaliza a posição do preço dentro das Bandas de Bollinger,

facilitando a interpretação da força do movimento dos preços.

O %B é definido pela seguinte fórmula:

$$\%B = \frac{P - BB_{\text{inferior}}}{BB_{\text{superior}} - BB_{\text{inferior}}}, \quad (2.16)$$

onde:

- P é o preço de fechamento do activo no período considerado;
- BB_{superior} é o valor da banda superior das Bandas de Bollinger;
- BB_{inferior} é o valor da banda inferior das Bandas de Bollinger.

O %B varia geralmente entre 0 e 1, indicando a posição relativa do preço entre as bandas inferior e superior. Valores próximos a 1 sugerem que o preço está próximo da banda superior, enquanto valores próximos a 0 indicam que o preço está próximo da banda inferior. Valores superiores a 1 ou inferiores a 0 indicam que o preço está fora das bandas, o que pode ser interpretado como sinais de sobrecompra ou sobrevenda, respectivamente (Blau, 2001).



Figura 2.4: Bandas de Bollinger e %B no EUR/USD.

Fonte: Metatrader5

Fluxo de dinheiro de Chaikin

O Fluxo de Dinheiro de Chaikin (FDC) é um indicador técnico utilizado para medir a pressão de compra e venda de um activo com base tanto no preço quanto no volume de negociação. Desenvolvido por Marc Chaikin, o indicador parte da premissa de que o volume pode antecipar movimentos de preço, ou seja, aumentos no volume durante altas indicam

força compradora, enquanto aumentos durante quedas sugerem pressão vendedora (Chaikin, 1992).

A lógica do FDC combina a posição do preço de fechamento dentro do intervalo diário com o volume negociado, sendo definido por:

$$FDC = \frac{\sum_{i=1}^n \left(\frac{(C_i - L_i) - (H_i - C_i)}{H_i - L_i} \times V_i \right)}{\sum_{i=1}^n V_i}, \quad (2.17)$$

onde:

- C_i : preço de fechamento no período i ;
- H_i : preço máximo no período i ;
- L_i : preço mínimo no período i ;
- V_i : volume negociado no período i ;
- n : número de períodos considerados (geralmente 20).

O valor desse indicador varia entre -1 e 1. Valores positivos indicam predominância da pressão compradora, enquanto valores negativos sugerem predominância da pressão vendedora.

O fluxo de dinheiro de Chaikin é considerado mais robusto do que outros indicadores baseados apenas em preço, pois incorpora o volume como variável essencial para a análise do comportamento do mercado (Chaikin, 1992).

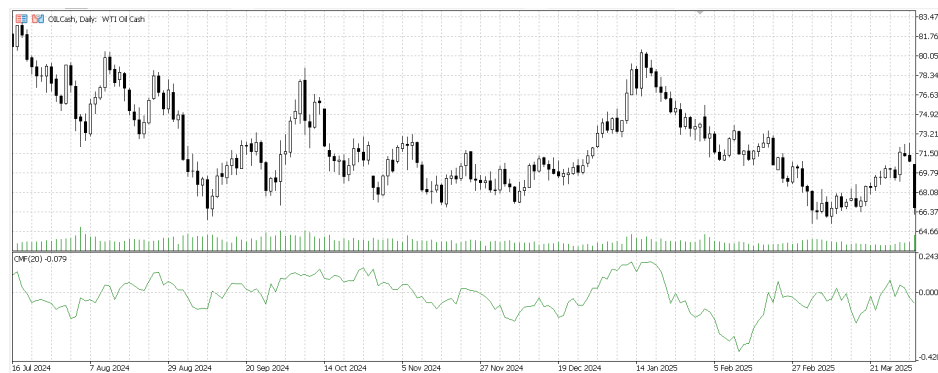


Figura 2.5: Fluxo de dinheiro de Chaikin do WTI.

Fonte: Metatrader5

Intervalo verdadeiro médio

O Intervalo Verdadeiro Médio, mais conhecido por ATR (*Average True Range*) é um indicador técnico desenvolvido por John Welles Wilder Jr. em 1978 com o objectivo de medir a volatilidade de um activo ao longo do tempo. Ao contrário de outros indicadores que se concentram na direcção do preço, o ATR é exclusivamente voltado para a intensidade dos movimentos de mercado, sendo amplamente utilizado por analistas técnicos para definir níveis de stop-loss e avaliar o risco operacional de um activo (Wilder, 1978).

A principal função do ATR é quantificar a volatilidade recente de um activo com base nas amplitudes de preço entre períodos consecutivos. O ATR é particularmente útil em mercados voláteis, pois leva em consideração os movimentos mais amplos, como oscilações entre máximas e mínimas.

O cálculo do ATR é baseado em duas etapas:

1. Intervalo Verdadeiro (TR):

Para cada período t , calcula-se o maior dos seguintes valores:

$$TR_t = \max \{H_t - L_t, |H_t - C_{t-1}|, |L_t - C_{t-1}|\} \quad (2.18)$$

Onde:

- H_t : preço máximo do período actual
- L_t : preço mínimo do período actual
- C_{t-1} : preço de fechamento do período anterior

2. Média do intervalo verdadeiro (ATR):

A média do intervalo verdadeiro é então suavizada por uma média móvel simples ou exponencial :

$$ATR_t = \frac{1}{n} \sum_{i=0}^{n-1} TR_{t-i}, \quad (2.19)$$

ou, no caso da média móvel exponencial (recomendada por Wilder, 1978):

$$ATR_t = ATR_{t-1} + \frac{1}{n}(TR_t - ATR_{t-1}), \quad (2.20)$$

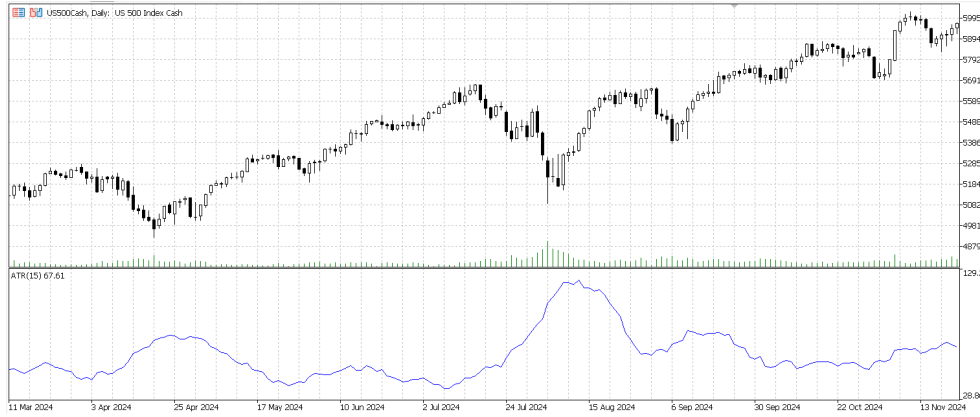


Figura 2.6: Intervalo verdadeiro médio do S&P500

Fonte: Metatrader5

Oscilador de momento de Chande

O oscilador de momento de chande (ou CMO, do inglês *Chande Momentum Oscillator*) é um indicador de momento. Ele foi projectado para medir a força do momentum de uma série de preços, avaliando directamente a diferença entre ganhos e perdas de uma série de preços ou retornos ao longo de um intervalo de tempo. Ao contrário de outros osciladores, como o IFR, o CMO considera valores absolutos de ganhos e perdas, tornando-se mais sensível a mudanças de direcção e capturando melhor a intensidade do movimento. Seu valor oscila no intervalo de $[-100, 100]$, permitindo identificar a dominância de altas ou baixas no período analisado (Chande, 1994).

A fórmula geral do CMO é dada por:

$$CMO_t = 100 \times \frac{Up_t - Down_t}{Up_t + Down_t}, \quad (2.21)$$

em que:

$$Up_t = \sum_{i=t-n+1}^t \max(r_i, 0) \quad \text{e} \quad Down_t = \sum_{i=t-n+1}^t \max(-r_i, 0), \quad (2.22)$$

sendo:

- r_i : variação no período i , que pode ser definida como diferença de preços ($r_i = P_i - P_{i-1}$);
- n : número de períodos da janela de análise;
- Up_t : soma das variações positivas (ganhos) nos últimos n períodos;
- $Down_t$: soma das variações negativas (perdas absolutas) nos últimos n períodos.

A interpretação do indicador é relativamente directa. Valores do CMO_t próximos de +100 indicam que os movimentos de alta predominam quase totalmente sobre os de baixa, reflectindo forte momentum positivo. De forma análoga, valores próximos de -100 indicam predominância de perdas e forte momentum negativo. Quando o indicador se aproxima de zero, significa que há relativo equilíbrio entre ganhos e perdas, sem uma tendência dominante clara. Em aplicações práticas, costuma-se adoptar os limiares de +50 e -50 como zonas de sobrecompra e sobrevenda, respectivamente, sinalizando possíveis reversões de tendência (Chande, 1994).

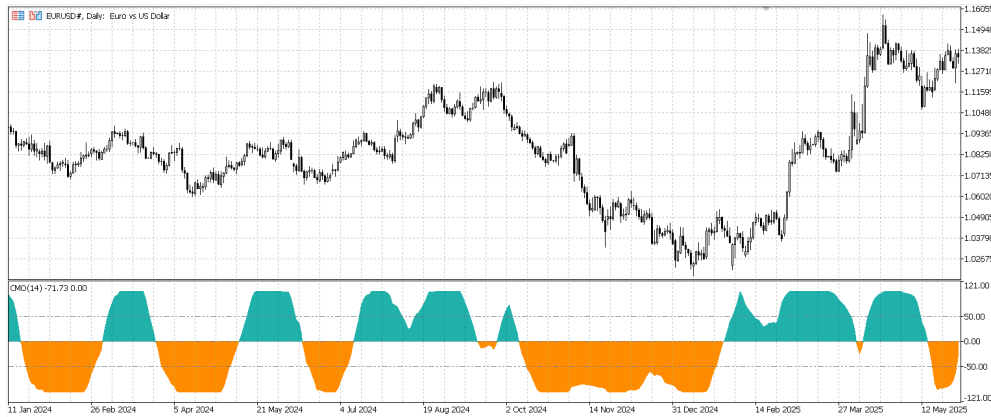


Figura 2.7: Oscilador de momento de Chande no EUR/USD

Fonte: Metatrader5

Oscilador estocástico

O oscilador estocástico é um indicador técnico desenvolvido por George Lane no final da década de 1950. Ele mede a posição relativa do preço de fechamento de um activo em

relação à faixa de preços mais alta e mais baixa dentro de um período específico. Os valores do oscilador variam de 0 a 100, indicando quão próximo o preço actual está do limite superior ou inferior da janela de observação (Lane, 1950).

O indicador é composto por dois componentes principais, %K e %D. O %K corresponde ao cálculo base, que localiza o preço de fechamento dentro da faixa de preços, enquanto o %D é uma versão suavizada da linha %K, funcionando como sinal de confirmação.

A formulação padrão é dada por:

$$\%K_t = \frac{C_t - L_t^{(p)}}{H_t^{(p)} - L_t^{(p)}} \times 100, \quad (2.23)$$

$$\%D_t = SMA(\%K_t, k), \quad (2.24)$$

onde:

- C_t é o preço de fechamento no período t ;
- p é o número de períodos considerados na janela de observação;
- $L_t^{(p)} = \min(C_{t-p+1}, \dots, C_t)$ é o menor preço observado nos últimos p períodos;
- $H_t^{(p)} = \max(C_{t-p+1}, \dots, C_t)$ é o maior preço observado nos últimos p períodos;
- $SMA(\%K_t, k)$ é a média móvel simples de k períodos da série %K, normalmente $k = 3$.

O %K indica a posição relativa do preço de fechamento dentro da faixa de preços recente, enquanto o %D suaviza essa informação para reduzir ruídos e fornecer sinais mais confiáveis.

A interpretação é baseada nos níveis de sobrecompra e sobrevenda. Valores acima de 80 sugerem que o activo está em sobrecompra, indicando possível reversão de tendência ou correcção de preços, por outro lado, valores abaixo de 20 indicam sobrevenda, sugerindo que o activo pode estar subavaliado e propenso a uma valorização (Lane, 1950).

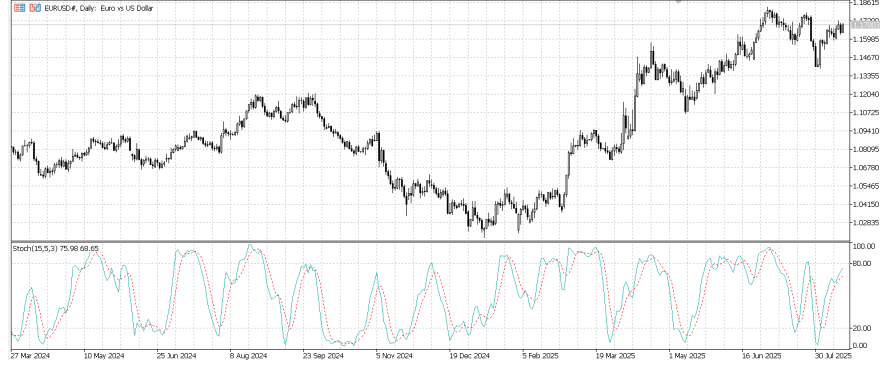


Figura 2.8: Oscilador Estocástico no EUR/USD.

Fonte: Metatrader5

Índice de movimento direccional

Do inglês *Directional Movement Index* (DMI), o Índice de Movimento Direccional, é um indicador técnico desenvolvido para medir a força e a direcção de uma tendência no mercado (Wilder, 1978). O DMI é composto por três elementos principais:

- +DI (indicador direccional positivo): mede a força dos movimentos ascendentes;
- −DI (indicador direccional negativo): mede a força dos movimentos descendentes;
- ADX (índice direccional medio): mede a força total da tendência, independentemente da direcção.

□ Cálculo do Movimento Direccional (DM):

$$DM_t^+ = \begin{cases} H_t - H_{t-1}, & \text{se } (H_t - H_{t-1} > L_{t-1} - L_t) \text{ e } H_t - H_{t-1} > 0 \\ 0, & \text{caso contrário} \end{cases}; \quad (2.25)$$

$$DM_t^- = \begin{cases} L_{t-1} - L_t, & \text{se } (L_{t-1} - L_t > H_t - H_{t-1}) \text{ e } L_{t-1} - L_t > 0 \\ 0, & \text{caso contrário} \end{cases}, \quad (2.26)$$

onde:

- H_t : preço máximo no período t ;
- L_t : preço mínimo no período t ;

- DM_t^+ : movimento direccional positivo;
- DM_t^- : movimento direccional negativo.

□ Cálculo do True Range (TR) (2.18) .

□ Cálculo dos indicadores $+DI$ e $-DI$:

$$+DI_t = 100 \times \frac{EMA(DM_t^+, n)}{EMA(TR_t, n)}; \quad (2.27)$$

$$-DI_t = 100 \times \frac{EMA(DM_t^-, n)}{EMA(TR_t, n)}, \quad (2.28)$$

onde $EMA(\cdot, n)$ é a média móvel exponencial de n períodos (tipicamente $n = 14$).

□ Cálculo do ADX:

$$DX_t = 100 \times \frac{|+DI_t - (-DI_t)|}{+DI_t + (-DI_t)}, \quad (2.29)$$

$$ADX_t = EMA(DX_t, n). \quad (2.30)$$

O ADX indica apenas a força da tendência, sem considerar a direcção, valores acima de 25 sugerem tendência forte, enquanto valores abaixo de 20 indicam tendência fraca ou mercado lateral. Quando $+DI > -DI$, a tendência predominante é de alta, enquanto que quando, $-DI > +DI$, indica que há tendência predominante é de baixa (Wilder, 1978).

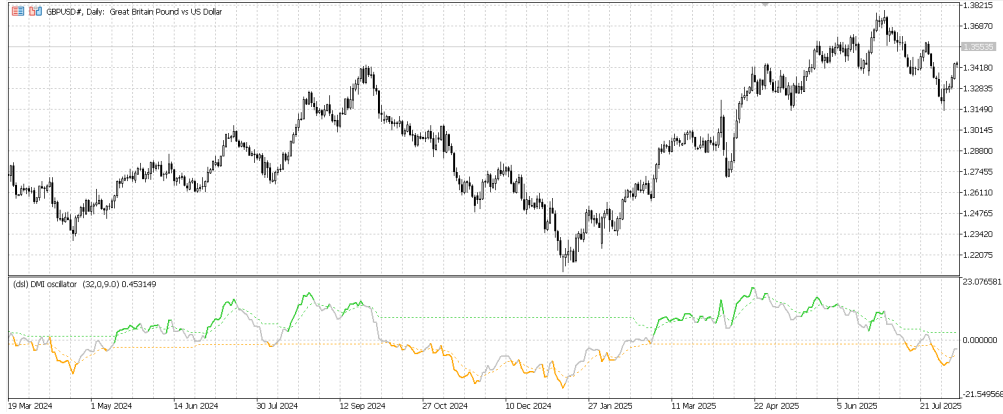


Figura 2.9: Índice de Movimento Direccional no GBP/USD.

Fonte: Metatrader5

Proporção do corpo da vela

A proporção do corpo da vela mede a intensidade do movimento direccional dentro de um período de negociação. Ele é definido como a razão entre o corpo real da vela, obtido pela diferença entre o preço de fechamento (C_t) e o preço de abertura (O_t), e a amplitude total da vela, calculada pela diferença entre o preço máximo (H_t) e o preço mínimo (L_t), e é expressa por:

$$Prop = \frac{|C_t - O_t|}{H_t - L_t}. \quad (2.31)$$

Valores próximos de 1 indicam que o período apresentou um movimento direccional forte, com predominância clara de compradores ou vendedores. Por outro lado, valores baixos sugerem indecisão no mercado, indicando que o preço permaneceu relativamente estável dentro do intervalo máximo-mínimo do período (Uzun et al., 2024).

2.5 Medidas estatísticas

As medidas estatísticas constituem instrumentos fundamentais para a descrição de dados, permitindo resumir características das distribuições e auxiliar na interpretação das informações observadas.

2.5.1 Medidas de forma da distribuição

As medidas de forma da distribuição descrevem características estruturais como assimetria e grau de achatamento, complementando a análise descritiva ao indicar desvios em relação

à simetria e à forma padrão de distribuições teóricas (Joanes & Gill, 1998).

Assimetria

A assimetria é uma medida estatística que indica o grau de simetria da distribuição das observações. Ela quantifica o quanto a distribuição das observações se desvia da simetria em torno da média (Joanes & Gill, 1998).

A fórmula da assimetria é dada por:

$$\text{Assimetria} = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^3, \quad (2.32)$$

onde:

- n é o número total de observações,
- X_i é o valor observado na i -ésima observação,
- $\bar{X}$ é a média das observações,
- σ é o desvio padrão das observações.

Nesta medida, valores positivos de assimetria indicam uma distribuição com cauda mais longa à direita, sugerindo maior probabilidade de observações extremas positivas. Já valores negativos indicam uma cauda mais longa à esquerda, evidenciando maior probabilidade de observações extremas negativas.

Curtose

A curtose é uma medida estatística que quantifica o grau de concentração dos dados nas caudas da distribuição das observações, indicando o peso das caudas em relação a uma distribuição normal.

A fórmula da curtose é dada por:

$$\text{Curtose} = \frac{1}{n} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^4, \quad (2.33)$$

onde:

- n é o número total de observações,
- X_i é o valor observado na i -ésima observação,
- $\bar{X}$ é a média dos valores,
- σ é o desvio padrão.

Valores de curtose muito próximos de 3 se aproximam de uma distribuição normal, enquanto valores de curtose maiores que 3 indicam distribuições com caudas mais pesadas e maior probabilidade de ocorrência de eventos extremos (leptocúrticas). Já valores menores que 3 sugerem distribuições mais achatadas (platicúrticas) em comparação à normal (Jondeau et al., 2007).

2.5.2 Medidas de variabilidade

As medidas de variabilidade quantificam o grau de dispersão dos dados em torno de uma tendência central como a média ou mediana. Entre essas métricas, destaca-se o coeficiente de variação, que expressa a variabilidade relativa em relação à média (Montgomery & Runger, 2014).

Coeficiente de variação

O Coeficiente de Variação (CV) é definido como a razão entre o desvio padrão e o valor absoluto da média de dados, sendo uma medida padronizada da dispersão relativa. Essa métrica permite avaliar a variabilidade de forma adimensional, possibilitando comparações entre conjuntos de dados com médias ou escalas diferentes (Fabozzi, 2007).

A sua formulação matemática é dada por:

$$CV = \frac{\sigma}{|\bar{X}|} * 100, \quad (2.34)$$

onde:

- σ é o desvio padrão dos, que mede a dispersão dos valores em torno da média (2.51);
- $\bar{X}$ é a média aritmética (2.36);

Por quantificar a dispersão de forma adimensional, o CV permite comparar a variabilidade entre diferentes amostras, mesmo quando apresentam escalas, magnitudes ou unidades distintas. Valores mais elevados indicam maior variabilidade relativa, ao passo que valores menores sugerem maior estabilidade dos dados.

Tabela 2.1: Classificação do coeficiente de variação

Classificação	Critério
Variação baixa	$0 < CV \leq 0,05$
Variação relativamente baixa	$0,05 < CV \leq 0,10$
Variação média	$0,10 < CV \leq 0,15$
Variação relativamente alta	$0,15 < CV \leq 0,20$
Variação alta	$CV > 0,20$

2.5.3 Medidas de dependência

As medidas de dependência descrevem o grau pelo qual duas variáveis se relacionam, permitindo quantificar padrões de associação presentes nos dados. Entre as abordagens utilizadas estão a:

Correlação de Pearson

A correlação de Pearson mede a relação linear entre duas variáveis X e Y . É calculada como:

$$\rho_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}}, \quad (2.35)$$

onde:

- n é o número total de observações;
- X_i é o valor da variável X na i -ésima observação;
- Y_i é o valor da variável Y na i -ésima observação;
- $\bar{X}$ é a média aritmética dos valores de X , dada por:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i; \quad (2.36)$$

- $\bar{Y}$ é a média aritmética dos valores de Y , dada por:

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i; \quad (2.37)$$

- O numerador $\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$ representa a covariância entre X e Y ;
- O denominador é o produto dos desvios padrão de X e Y , dados por:

$$\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}. \quad (2.38)$$

Uma correlação positiva ($\rho_{XY} > 0$) sugere que as variáveis tendem a mover-se na mesma direcção, enquanto uma correlação negativa ($\rho_{XY} < 0$) indica que se movem em direcção oposta. Valores de ρ_{XY} próximos de zero indicam fraca ou inexistente relação linear, pelo que, valores próximos de +1 indicam forte relação linear crescente, enquanto valores próximos de -1 indicam forte relação linear decrescente. (Tsay, 2010).

Correlação de Spearman

A correlação de Spearman mede a relação monótona entre duas variáveis X e Y . Em vez de considerar os valores originais, utiliza os postos (*ranks*) de cada observação. É calculada como:

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (2.39)$$

onde:

- n é o número total de observações;
- $d_i = R(X_i) - R(Y_i)$ é a diferença entre os postos da i -ésima observação de X e Y ;
- $R(X_i)$ é o posto (ordem crescente) associado ao valor X_i ;
- $R(Y_i)$ é o posto (ordem crescente) associado ao valor Y_i .

A interpretação da correlação de Spearman é idêntica a da correlação de Pearson (Conover, 1999).

2.6 Séries temporais

As séries temporais são sequências de observações ao longo do tempo, nas quais a ordem temporal influencia directamente as propriedades estatísticas do processo. Diferentemente de dados independentes, séries temporais podem apresentar dependência entre observações, tendência, sazonalidade e estruturas de autocorrelação (Hamilton, 1994; Chatfield, 2004).

Embora as séries temporais sejam aplicáveis em diversos domínios, desde ciências naturais até engenharia, destacam-se também em aplicações financeiras, onde retornos, volatilidade e índices exibem padrões temporais que motivam a análises destas (Tsay, 2010).

2.6.1 Operações sobre séries temporais

Defasagem

O defasamento (ou *lag*) é uma operação aplicada a séries temporais que consiste em deslocar os valores de uma variável para períodos anteriores no tempo, criando versões atrasadas dessa variável. Em séries temporais financeiras, essa técnica é amplamente utilizada para capturar relações temporais, dependência entre observações passadas e actuais, e para criar variáveis explicativas baseadas em informações históricas (Box, 2015; Tsay, 2010).

A formulação geral para uma defasagem de ordem k é:

$$X_{t-k}, \tag{2.40}$$

onde:

- X_t representa o valor da variável no período t ;
- k é a ordem da defasagem, isto é, o número de períodos deslocados;
- X_{t-k} representa o valor da variável X k períodos antes do tempo t .

Suavização

Técnicas como médias móveis simples e exponenciais são comumente aplicadas para reduzir ruído e destacar tendências. Essas técnicas são detalhadas na secção 2.4.

2.6.2 Propriedades de séries temporais

Estacionariedade e teste de Dickey-Fuller aumentado

Uma série temporal é dita estacionária quando suas propriedades estatísticas, como média, variância e autocorrelação, permanecem constantes ao longo do tempo (Hamilton, 1994).

Segundo Morettin e Toloi (2011), uma série temporal é estacionária se, ela evolui aleatoriamente em torno de uma média constante ao longo do tempo, reflectindo uma certa de estabilidade de equilíbrio.

O Teste de Dickey-Fuller Aumentado (ADF) é amplamente utilizado para avaliar a presença de uma raiz unitária, indicativa de não estacionariedade (Dickey, 1979). Sua formulação geral é dada por:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t, \quad (2.41)$$

em que:

- y_t é o valor da série temporal no tempo t ;
- $\Delta y_t = y_t - y_{t-1}$ é a primeira diferença da série;
- α é o intercepto;
- βt é um termo de tendência determinística;
- γ é o coeficiente da variável defasada, utilizado para testar a presença de raiz unitária;
- p é o número de defasagens adicionais para corrigir autocorrelação serial;
- ε_t é o termo de erro aleatório.

O teste ADF é baseado nas seguintes hipóteses:

- H_0 : a série apresenta raiz unitária, isto é, não estacionária;
- H_a : a série não apresenta raiz unitária, sendo assim estacionária.

O critério de decisão do teste ADF baseia-se na comparação entre o valor- p e o nível de significância α (por exemplo, 0,05). Se o valor- p for menor que α , rejeita-se a hipótese nula (H_0) e conclui-se que a série é estacionária. Por outro lado, se o valor- p for maior ou igual a α , não se rejeita H_0 , indicando que a série não é estacionária. O teste ADF é fundamental na análise de séries temporais, pois ajuda a evitar problemas como regressões espúrias, que podem ocorrer quando variáveis não estacionárias são utilizadas sem diferenciação (Enders, 2014).

Coeficiente de Hurst

O coeficiente de Hurst fornece uma medida da memória de longo prazo de uma série temporal. Como é robusto e requer poucas suposições, possui ampla aplicabilidade na análise de séries temporais. Os valores do expoente de Hurst variam entre 0 e 1. Com base no valor do expoente de Hurst (H), uma série temporal pode ser classificada em três categorias. $H = 0,5$, que indica uma série aleatória (sem memória). De acordo com Qian e Rasheed (2004), o intervalo $0 < H < 0,5$, indica uma série antipersistente (série com reversão à média) enquanto que $0,5 < H < 1$ indica uma série com comportamento persistente.

Uma série antipersistente possui a característica de reversão à média, o que significa que um valor alto é mais provável de ser seguido por um valor baixo, e vice-versa. A força dessa reversão à média aumenta à medida que H se aproxima de 0. Por outro lado, uma série persistente reforça a tendência, o que significa que a direcção do próximo valor é mais provável de ser a mesma que a do valor actual. A força da tendência aumenta à medida que H se aproxima de 1 (Qian & Rasheed, 2004).

Estudos demonstram que mercados com expoente de Hurst (H) superior a 0,5 (indicando menor eficiência), apresentam maior capacidade de previsão. Além disso, mercados emergentes tendem a exibir valores de H mais elevados e maior previsibilidade comparados a mercados desenvolvidos (Qian & Rasheed, 2004).

O expoente de Hurst pode ser calculado por análise de amplitude redimensionada (análise R/S). Para uma série temporal, $X = X_1, X_2, \dots, X_n$, o método de análise R/S é o seguinte:

1. Cálculo da média

$$m = \frac{1}{n} \sum_{i=1}^n X_i, \quad (2.42)$$

2. Cálculo da série ajustada Y

$$Y_t = X_t - m, \quad t = 1, 2, \dots, n, \quad (2.43)$$

3. Cálculo do desvio cumulativo Z

$$Z_t = \sum_{i=1}^t Y_i, \quad t = 1, 2, \dots, n, \quad (2.44)$$

4. Cálculo da série de amplitude R

$$R(n) = \max(Z_1, Z_2, \dots, Z_n) - \min(Z_1, Z_2, \dots, Z_n), \quad (2.45)$$

5. Cálculo da série de desvio padrão S

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - m)^2}, \quad (2.46)$$

onde m é a média calculada no ponto 1.

6. Cálculo do alcance reescalado

$$\left(\frac{R}{S}\right)_n = \frac{R(n)}{S(n)}. \quad (2.47)$$

Sendo que $\left(\frac{R}{S}\right)_n$ segue uma lei de potência quando o tempo aumenta, resultando na equação:

$$E\left(\frac{R}{S}\right)_n = c n^H, \quad (2.48)$$

onde, c é uma constante independente de n e H é o expoente de Hurst.

Para estimar H , aplica-se uma regressão linear simples aos logaritmos naturais da equação:

$$\ln E\left(\frac{R}{S}\right)_n = \ln c + H \ln n. \quad (2.49)$$

de tal forma que o H será o coeficiente angular da regressão linear simples.

2.7 Estimadores de volatilidade

No contexto deste estudo, a volatilidade pode ser entendida, em termos estatísticos, como a variabilidade de uma série temporal ao longo do tempo, sendo normalmente associada à

variância ou ao desvio padrão do processo (Tsay, 2010). A literatura apresenta diversos estimadores de volatilidade, com diferentes pressupostos e propriedades estatísticas. Esses estimadores constituem alternativas amplamente discutidas na literatura estatística e econométrica para a avaliação da variabilidade de séries temporais. Para fins de revisão e aplicação neste estudo, destacam-se:

Volatilidade baseada no desvio padrão anualizado

A volatilidade baseada no desvio padrão anualizado é uma medida estatística que quantifica a dispersão dos retornos de um activo financeiro ao longo do tempo, ajustando essa variabilidade para uma base anual. É amplamente utilizada na gestão de riscos e na avaliação da incerteza associada ao retorno de activos, especialmente em contextos como precificação de derivativos, estratégias de cobertura e modelos de alocação de portfólio (Hull, 2015).

Sua função principal é indicar o risco esperado de um activo em termos de dispersão de retornos. Quanto maior a volatilidade, maior o risco percebido, e, por consequência, maior pode ser o prémio exigido pelos investidores para compensar essa incerteza. A volatilidade anualizada também é um insumo fundamental em modelos como o CAPM e o modelo de precificação de opções Black e Scholes (Hull, 2015).

A fórmula padrão para a volatilidade anualizada dos retornos é:

$$\sigma_{\text{anualizada}} = \sigma \times \sqrt{N} \quad (2.50)$$

Sendo:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}, \quad (2.51)$$

Onde, σ é o desvio padrão dos retornos; N é o número de períodos de negociação por ano (geralmente 252 dias); $\bar{X}$ é a média dos retornos e X_i é a observação i .

Na prática, a anualização permite comparar o risco de diferentes activos sob uma mesma escala temporal, facilitando análises comparativas e decisões de portfólio.

Volatilidade baseada no intervalo

As medidas de volatilidade baseadas no intervalo de preços, figuram entre os estimadores mais eficientes e informativos para capturar a variabilidade durante o período em análise. Segundo Gupta e Mishra (2024), essa classe de métodos utiliza a amplitude entre o preço máximo (H_t) e o preço mínimo (L_t) como fonte primária de informação, explorando o facto de que o intervalo contém muito mais sinal sobre a dispersão dos preços do que estimativas derivadas apenas de retornos de fechamento. Sua formulação geral é dada por:

$$R_t = \frac{(H_t - L_t)}{L_t}, \quad (2.52)$$

onde R_t representa a amplitude proporcional do intervalo de negociação no período t . Este estimador é simples, facilmente computável e, conforme destacado por Gupta e Mishra (2024), fornece uma medida directa da intensidade das oscilações ao longo do dia de negociação, tornando-se especialmente útil em ambientes com alta incerteza ou pressões de compra e venda intensificadas.

Volatilidade de Rogers e Satchell

A volatilidade de Rogers e Satchell é um estimador de volatilidade que assume retorno médio não nulo. Proposto por Rogers e Satchell (1991), este estimador é considerado não viesado mesmo na presença de tendência (*drift*) nos preços, o que o torna particularmente adequado para mercados reais.

Esse estimador utiliza os preços de abertura, fechamento, máxima e mínima para capturar a variabilidade intradiária dos preços. Sua principal vantagem é a robustez diante de processos estocásticos com tendência, permitindo uma estimativa eficiente da variância condicional. A fórmula da volatilidade de Rogers e Satchell para um período t , é dada por:

$$RS_t = [(\ln H_{i,t} - \ln O_{i,t}) \cdot (\ln H_{i,t} - \ln C_{i,t}) + (\ln L_{i,t} - \ln O_{i,t}) \cdot (\ln L_{i,t} - \ln C_{i,t})] \quad (2.53)$$

Onde: $H_{i,t}$, $L_{i,t}$, $O_{i,t}$ e $C_{i,t}$ são, respectivamente, os preços máximos, mínimos, de abertura e de fechamento do intervalo i do período t .

O estimador de volatilidade de Rogers e Satchell apresenta desempenho superior ao de estimadores baseados apenas em retorno de fechamento, especialmente em dados de alta

frequência (Ślepaczuk & Zakrzewski, 2009; Rogers & Satchell, 1991).

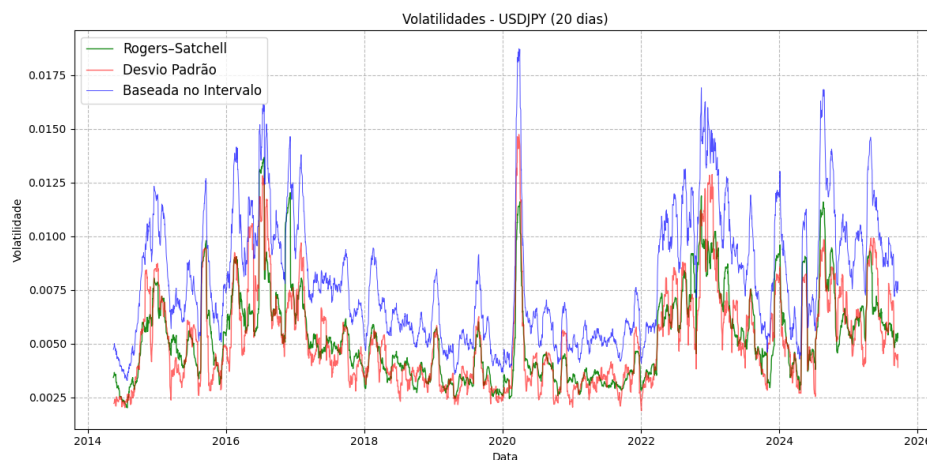


Figura 2.10: Comparação entre volatilidades no par USD/JPY

Fonte: Autor

2.8 Medidas de risco e desempenho financeiro

As medidas de risco e desempenho financeiro são ferramentas essenciais na análise de estratégias financeiras, pois permitem quantificar tanto os retornos esperados quanto a incerteza associada e a magnitude das perdas potenciais (Farinha Rodrigues, 2019). Essas medidas fornecem uma base para avaliar a robustez das estratégias e comparar abordagens diferentes, oferecendo uma visão mais completa da relação risco-retorno.

Enquanto as medidas de risco avaliam a exposição a eventos adversos e a volatilidade dos retornos, as medidas de desempenho financeiro permitem relacionar esses riscos ao retorno gerado, oferecendo indicadores ajustados ao risco que facilitam a comparação entre diferentes estratégias. Essa distinção é fundamental para uma análise financeira mais profunda, permitindo entender não apenas se uma estratégia gera lucro, mas também quão eficiente ela é diante do risco assumido (Rego, 2014).

No contexto do mercado Forex, caracterizado por alta volatilidade, essas métricas tornam-se ainda mais relevantes. Estudos de negociação algorítmica e previsão de pares de moedas frequentemente utilizam medidas como valor em risco, perda máxima acumulada, índice de Sharpe e índice de Sortino para avaliar a eficácia das estratégias e garantir que o risco de perdas extremas seja controlado (Zhang & Khushi, 2020).

Tanto para as medidas de risco quanto para as de desempenho financeiro, o conceito de retorno é uma das principais medidas utilizadas. Segundo Bodie, Kane e Marcus (2014),

o retorno de um investimento consiste na variação do valor do activo somada a qualquer rendimento em dinheiro (como dividendos ou juros) recebido durante o período de detenção. Essa definição converge com a visão de Morettin (2011), que descreve o retorno como o ganho relativo ou absoluto resultante da diferença entre o capital disponível no futuro e o montante aplicado no presente.

O retorno pode ser mensurado de duas formas principais: absoluta e relativa. O retorno absoluto é a expressão monetária do ganho, definida pela diferença directa entre o valor final e o valor inicial investido. No entanto, para fins de comparação entre diferentes activos e períodos, a taxa de retorno (ou retorno relativo) é a medida mais utilizada. Sendo expressa pela Equação 2.54:

$$\text{Taxa de Retorno} = \frac{\text{valor corrente} - \text{valor inicial investido}}{\text{valor inicial investido}}. \quad (2.54)$$

Medidas de risco

- **Value at Risk (VaR)**

O Valor em risco (*Value at Risk*) mede a maior perda percentual, com um determinado nível de confiança, que se espera que possa ocorrer em uma posição de portfólio (ou activo) ao longo de um dado período de detenção (Jorion, 2007). Matematicamente, para uma variável aleatória de retornos R com função distribuição $F_R(r)$, o VaR pode ser definido como o quantil da cauda inferior, expresso como:

$$VaR_\alpha = -\inf\{r \in \mathbb{R} : F_R(r) \geq 1 - \alpha\}, \quad (2.55)$$

ou, de forma equivalente,

$$VaR_\alpha = -Q_R(1 - \alpha), \quad (2.56)$$

onde $F_R(r) = P(R \leq r)$ corresponde à função de distribuição acumulada desses retornos; α denota o nível de significância, indicando a probabilidade de que a perda exceda o VaR; e $Q_R(1-\alpha)$ é o quantil inferior de ordem $1-\alpha$ da distribuição de R , que determina directamente o valor do VaR sob uma abordagem não paramétrica.

Por outro lado, assumindo que os retornos seguem uma distribuição normal, $R \sim N(\mu, \sigma^2)$,

isto é, adoptando a abordagem paramétrica, o VaR pode ser obtido pela padronização da variável. Definindo $X = R$, tem-se:

$$\begin{aligned}
 P(X \leq VaR_\alpha) &= 1 - \alpha, \\
 P\left(\frac{X - \mu}{\sigma} \leq \frac{VaR_\alpha - \mu}{\sigma}\right) &= 1 - \alpha, \\
 P(Z \leq z_{1-\alpha}) &= 1 - \alpha, \\
 VaR_\alpha &= \mu + z_{1-\alpha}\sigma.
 \end{aligned} \tag{2.57}$$

onde:

- μ é a média dos retornos;
- σ é o desvio padrão dos retornos;
- Z é uma variável aleatória com distribuição normal padrão;
- $z_{1-\alpha} = \Phi^{-1}(1 - \alpha)$ é o quantil da normal padrão correspondente à probabilidade $1 - \alpha$;
- $\Phi(\cdot)$ é a função distribuição acumulada da normal padrão: $\Phi(z) = P(Z \leq z)$.

Nas expressões apresentadas anteriormente para o VaR paramétrico (2.57) e não paramétrico (2.56), é possível classificar os métodos de estimação em diferentes abordagens. A abordagem paramétrica, amplamente discutida por Jorion (2007), assume uma distribuição específica para os retornos, tipicamente a normal. Já métodos não paramétricos, como a simulação histórica, utilizam o quantil empírico sem assumir hipóteses de distribuição (Dowd, 2002).

Pesquisas recentes indicam que abordagens baseadas em quantis empíricos apresentam melhor desempenho para capturar riscos extremos (Guo et al., 2023). Além disso, estudos aplicados ao mercado Forex, como os de Bredin e Hyde (2018) e Cai (2007), demonstram que o VaR pode ser utilizado como uma ferramenta de gestão de risco, servindo de referencial para a estipulação de limites de perda em operações e estratégias de trading.

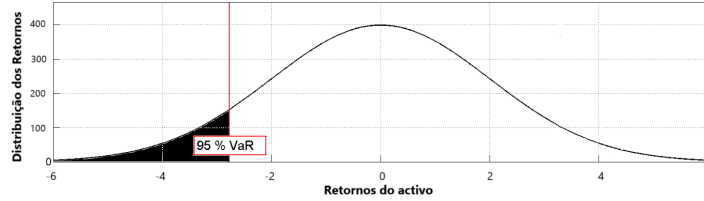


Figura 2.11: Ilustração do VaR de $\alpha = 5\%$ de uma distribuição normal de retornos.

Fonte: Adaptado de Kisiala, J. (2015).

- **Conditional Value at Risk (CVaR)**

O Valor em Risco Condicional (CVaR) é uma extensão do VaR que busca superar suas limitações, principalmente a incapacidade de capturar a magnitude das perdas extremas. Enquanto o VaR identifica apenas o quantil de perda com probabilidade α , o CVaR considera a média das perdas que excedem esse quantil, oferecendo uma visão mais completa do risco de cauda (Rockafellar, 2000). Formalmente, para uma variável de retornos R com função de distribuição $F_R(r)$:

$$CVaR_\alpha = \mathbb{E}[-R \mid R \leq -VaR_\alpha] .$$

Sua relevância na literatura financeira decorre do facto de capturar eventos extremos e permitir optimização de portfólio (ou retorno do activo) robusta, sobretudo em mercados com caudas pesadas e volatilidade elevada (Rockafellar, 2000; Guo, 2023).

Medidas de desempenho

- **Índice Sharpe**

O Índice de Sharpe (IS) é a medida padrão e fundamental para avaliar o desempenho de um investimento ajustado ao risco, introduzida por William F. Sharpe em 1966. O objectivo principal deste índice é quantificar a recompensa do risco, ou seja, o retorno excedente que uma estratégia de investimento entrega por unidade de volatilidade total assumida.

O cálculo é dado por:

$$IS = \frac{\mu_R - r_f}{\sigma_R},$$

Onde μ_R é o retorno médio do portfólio, activo ou da estratégia de investimento; r_f é a taxa de retorno do activo livre de risco e σ_R é o desvio padrão ou volatilidade dos retornos do portfólio, activo ou da estratégia de investimento.

De acordo com Júnior (2009) o índice de Sharpe é o melhor indicador para avaliar o desempenho de risco sobre o retorno, além disso, o efeito do Sharpe acaba sendo maior quando o risco da empresa é menor, portanto, apresentaria o maior indicador e o melhor desempenho do retorno sobre o risco.

Estudos aplicados ao mercado de câmbio, particularmente em análises de estratégias de trading, a taxa livre de risco (r_f) é usualmente considerada nula (Lo, 2002).

• Índice Sortino

O índice sortino foi desenvolvido como uma alternativa ao Sharpe, buscando medir desempenho ajustado ao risco de forma mais consistente com a aversão a perdas. Enquanto o Sharpe penaliza toda a volatilidade, o Sortino considera apenas a variação dos retornos que ficam abaixo de um retorno mínimo aceitável (MAR). Essa abordagem concentra-se no conceito de risco de queda (*downside risk*), reflectindo melhor o comportamento real dos investidores diante do risco (Sortino, 1991).

$$Sortino = \frac{\mu_R - MAR}{\sigma_-},$$

onde μ_R é o retorno médio; MAR é o retorno mínimo aceitável definido pelo investidor; σ_- é o desvio padrão negativo, que representa o risco associado apenas às perdas, e é dado por:

$$\sigma_- = \sqrt{\frac{1}{T} \sum_{t=1}^T \min(0, R_t - MAR)^2}.$$

Sendo que T representa o número total de observações; R_t é o retorno no período t ; $\min(0, R_t - MAR)$ captura apenas os desvios negativos em relação ao MAR.

Ao ignorar a volatilidade positiva, o Sortino evita penalizar flutuações benéficas e torna-se especialmente adequado para distribuições de retornos assimétricas, estratégias com assimetria estrutural ou quando há ocorrência de caudas pesadas. A literatura evidencia que essa medida fornece uma avaliação mais realista do desempenho de fundos e estratégias quantitativas em ambientes não simétricos de risco (Sortino & Van der Meer, 1994).

• Drawdown e maximum drawdown

O drawdown é uma medida amplamente utilizada na gestão de risco para quantificar perdas

em activos, investimentos ou portfólios. Ele representa a queda percentual do valor de um activo ou portfólio entre um pico local e o subsequente vale, permitindo avaliar a severidade de perdas temporárias em determinado período (Geronazzo, 2019). Formalmente, o drawdown no período t pode ser definido como:

$$DD_t = \frac{V_t - V_{\max,t}}{V_{\max,t}},$$

onde V_t é o valor do activo ou portfólio no tempo t ; $V_{\max,t} = \max(V_1, V_2, \dots, V_t)$ é o valor máximo observado até o período t ; Quando o $DD_t \leq 0$ indica perda relativa em relação ao pico anterior.

O *Maximum Drawdown* (MDD) corresponde à maior queda observada ao longo de toda a trajectória do valor de uma estratégia de investimento ou portfólio, representando o risco extremo enfrentado pelo investidor ou trader (Goldberg e Mahmoud, 2014; Atiya & Magdon-Ismail, 2018).

$$MDD = \min(DD_1, DD_2, \dots, DD_T),$$

com T sendo o número total de períodos analisados. Embora cada trajectória apresente apenas um MDD, a análise da distribuição desta medida em múltiplos períodos históricos permite avaliar a frequência e a profundidade de perdas potenciais, auxiliando na tomada de decisões de gestão de risco.

Na prática, o MDD é amplamente utilizado como referência para estabelecer limites de stop-loss, ajustar a alocação de capital e calibrar estratégias de gestão de risco, uma vez que fornece uma estimativa do pior cenário de perda que um investimento pode enfrentar (Chekhlov, Uryasev & Zabarankin, 2005; Goldberg & Mahmoud, 2016).

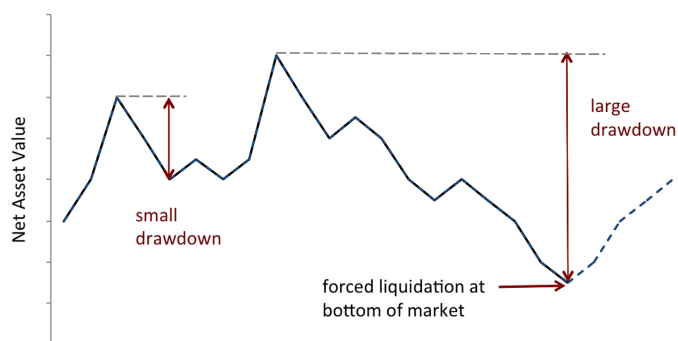


Figura 2.12: Ilustração de drawdowns no valor de um portfólio ao longo de um período de tempo

Fonte: Goldberg e Mahmoud (2016).

2.9 Machine learning

Machine Learning ou aprendizado de máquina pode ser entendida como um conjunto de métodos estatísticos e computacionais destinados a identificar padrões em dados e produzir modelos ou algoritmos capazes de realizar previsões ou decisões baseadas em tais padrões. Segundo Hastie, Tibshirani e Friedman (2009), o objectivo fundamental consiste em desenvolver algoritmos que utilizem um conjunto de observações para inferir uma função de forma aproximada que descreve as observações. Bishop (2006) reforça que esses métodos buscam generalizar a partir de dados observados, minimizando o erro de previsão em novas observações não vistas.

2.9.1 Aprendizagem supervisionada

A aprendizagem de máquina pode ser dividida em diferentes grupos, de acordo com a disponibilidade e o tipo de informação sobre os dados. Um dos grupos mais utilizados é a aprendizagem supervisionada, que se caracteriza pelo uso de dados rotulados para treinar os modelos. No contexto da aprendizagem supervisionada, o algoritmo recebe um conjunto de pares entrada e saída, com o objectivo de aprender uma função que mapeie as entradas às saídas desejadas (Bishop, 2006).

De acordo com Goodfellow, Bengio e Courville (2016), o termo supervisionada decorre do facto de que o processo de aprendizagem é conduzido por variável resposta (*alvo*) previamente conhecidos, permitindo ao modelo ajustar seus parâmetros para minimizar uma função de perda que quantifica o erro entre a previsão e o valor verdadeiro.

Um ponto conceitual essencial na aprendizagem supervisionada é a distinção entre problemas de classificação e regressão. Em problemas de classificação, o objectivo é atribuir cada observação a uma categoria previamente definida, sendo a variável resposta qualitativa ou discreta. Trabalhos clássicos como Duda et al. (2001) definem classificação como o processo de atribuir vectores de características a um conjunto finito de classes mutuamente exclusivas.

Por outro lado, em problemas de regressão, o alvo é quantitativo e o objectivo consiste em modelar uma função que estime valores contínuos. Hastie et al. (2009) destaca que ambos os tipos de problemas são abordados sob a mesma estrutura estatística geral, aprender uma relação funcional que minimize o erro em dados futuros, preservando a capacidade de generalização.

2.9.2 Etapas de abordagem de machine learning

Estas etapas incluem a recolha dos dados, a análise exploratória, a preparação dos conjuntos de treino e validação, o treinamento dos algoritmos e a sua posterior optimização.

A abordagem de machine learning envolve um conjunto estruturado de etapas que visam garantir a qualidade dos dados, a robustez dos modelos e a validade dos resultados. De acordo com Hastie et al. (2021), o desempenho de modelos predictivos depende fortemente da forma como os dados são preparados, analisados, divididos e utilizados no processo de treino e validação.

2.9.2.1 Colecção de dados

A colecção de dados refere-se à obtenção e organização das observações que serão utilizadas no desenvolvimento de algoritmos de machine learning.

Esta é a etapa fundamental em que os dados brutos são colectados a partir de diversas fontes, como logs, bancos de dados ou APIs externas. A colecta de dados de alta qualidade garante que o modelo de aprendizado de máquina seja treinado com conjuntos de dados relevantes e representativos (Yuji Roh et al., 2019).

Provost e Fawcett (2013) também destacam, que a qualidade dos dados influencia significativamente a precisão dos modelos, sendo essencial garantir que os dados sejam representativos do fenómeno estudado e livres de erros estruturais ou lógicos.

Em contextos de séries temporais, a preservação da ordem cronológica e a definição adequada do intervalo temporal de análise são aspetos essenciais. A integridade da sequência temporal assegura que os algoritmos aprendam padrões representativos sem introdução de informação futura (*data leakage*), o que poderia enviesar as estimativas de desempenho (Bergmeir & Benítez, 2012).

2.9.2.2 Análise exploratória de dados

A Análise Exploratória de Dados (EDA) é uma etapa preliminar essencial em qualquer fluxo de trabalho em machine learning. Ela visa caracterizar a estrutura, as propriedades estatísticas e as relações entre variáveis no conjunto de dados antes da construção dos modelos. Segundo Tukey (1977), pioneiro no conceito de EDA, esta análise permite revelar padrões, detectar anomalias e gerar hipóteses que podem orientar as etapas subsequentes de modelação. Algumas das ferramentas mais usadas na EAD são a matriz de correlação, o grafico de dispersão, o boxplot, o histograma e os gráficos de barras.

Hyndman e Athanasopoulos (2018) enfatizam que a EDA podem incluir a avaliação de características como heterocedasticidade e dependências temporais. A identificação de tais atributos é fundamental para determinar a necessidade de transformações (por exemplo, diferenciação ou normalização) que podem melhorar a capacidade de aprendizagem dos modelos.

A identificação e análise de valores extremos e valores faltantes também é crítica e é efectuada durante o processo de EAD. Little e Rubin (2019) apontam que dados ausentes ou extremos podem introduzir vieses significativos se não forem tratados adequadamente, comprometendo a validade das medidas de desempenho dos modelos.

2.9.2.3 Técnicas de pré-processamento de dados

O pré-processamento de dados consiste em um conjunto de técnicas utilizadas para preparar e organizar os dados brutos antes de sua análise ou aplicação em modelos. Essa etapa é fundamental para garantir a qualidade e a consistência das informações, permitindo que os algoritmos de machine learning operem de forma mais eficiente e confiável. O pré-processamento pode envolver a limpeza, transformação dos dados, tratamento de valores ausentes e outliers (Kuhn & Johnson, 2013).

Centramento e escalonamento: O centramento e o escalonamento são transformações fundamentais para garantir que todos os predictores estejam na mesma escala, facilitando a convergência e a estabilidade numérica de diversos modelos. Segundo Kuhn e Johnson (2013), centrar consiste em subtrair a média da variável, enquanto escalar divide seus valores pelo desvio padrão, resultando em predictores com média zero e desvio padrão igual a um. A principal limitação do centramento e do escalonamento, está na perda de interpretabilidade das unidades originais.

Tratamento de valores extremos: Valores extremos (*outliers*) podem afectar estimativas, ampliar variâncias e prejudicar o desempenho dos modelos. Para identificá-los, técnicas clássicas como Z-score ou box-plot são úteis, existem porém métodos modernos como o floresta de isolamento (Liu et al., 2008), que oferecem detecção automatizada de anomalias. No seu tratamento os valores extremos podem ser removidos ou transformados (Atkinson, 2021).

Tratamento de valores ausentes: Valores faltantes podem ser excluídos ou imputados por métodos simples, como média, mediana ou moda, ou por técnicas mais avançadas, como regressão (Van, 2018).

2.9.2.4 Divisão do conjunto de dados

A divisão do conjunto de dados em subconjuntos de treino e de validação (ou teste) é uma etapa essencial em machine learning, pois permite estimar a capacidade de generalização dos modelos em dados não observados durante o treino.

Em problemas tradicionais de machine learning, a divisão aleatória em amostras de treino e teste é frequentemente adoptada (Kohavi, 1995). Contudo, quando se lida com séries

temporais, a simples amostragem aleatória pode introduzir informação futura indevida no conjunto de treino, resultando em estimativas de desempenho excessivamente optimistas (Bergmeir & Benítez, 2012).

De acordo com Hyndman e Athanasopoulos (2018), a divisão temporal é a abordagem adequada para séries temporais, primeiro usa-se um período inicial para treino, seguido por um período mais recente para validação. Esta estratégia preserva a ordem cronológica dos eventos, evitando a introdução de vazamento de informação (*data leakage*), e simula as condições de previsão em ambiente real.

Uma forma comum de divisão temporal é a separação em blocos definidos, por exemplo, 70% dos dados usados para treino e os restantes 30% para validação ou teste, respeitando sempre a sequência cronológica (Bergmeir & Benítez, 2012).

2.9.2.5 Discretização de variáveis

A discretização, conhecida em inglês como *binning*, é um processo de transformação de variáveis contínuas em categorias, agrupando valores em bins ou intervalos com o objectivo de reduzir ruído, lidar com não linearidades ou melhorar a interpretabilidade de modelos. Segundo Kotsiantis e Kanellopoulos (2006), a discretização pode aumentar a estabilidade dos algoritmos. Liu et al. (2002), destacam que bons métodos de discretização podem preservar a informação relevante reduzindo a variância do modelo. A literatura divide esse processo em métodos não supervisionados, que não utilizam o rótulo y , e supervisionados, que consideram a relação da variável com o y .

Alguns dos métodos de discretização não supervisionados são:

- **Discretização por largura igual:** Este método, consiste em dividir o intervalo total da variável contínua em partes de mesma largura. Conforme descrevem Han, Kamber e Pei (2012), esse método não considera a distribuição dos dados, o que pode resultar em *bins* com poucas observações em regiões onde a densidade é baixa e bins excessivamente preenchidos onde a densidade é alta. Apesar da sua simplicidade, em modelos sensíveis a valores extremos, este pode sofrer distorções, pois valores extremos expandem artificialmente o intervalo dos *bins*.
- **Discretização por quantis:** É um método que divide os dados de forma que cada *bin* contenha aproximadamente o mesmo número de observações. Segundo Dougherty, Kohavi e Sahami (1995), esse método costuma ser mais robusto à presença de assimetria e concentrações de dados, pois garante equilíbrio entre *bins* independentemente da forma da distribuição. Entretanto, o método pode gerar intervalos estreitos em regiões

de alta densidade, o que pode suavizar estruturas importantes dos dados e aumentar o risco de sobreajuste em modelos sensíveis a limites artificiais.

Por outro lado a discretização pode ser supervisionada, e alguns dos métodos são:

- **Discretização por análise bivariada** - Este método consiste em definir pontos de corte avaliando directamente a relação entre cada variável predictor e o alvo. Segundo Siddiqi (2006), este tipo de discretização é guiada por critérios estatísticos que buscam maximizar a força e a estabilidade da associação entre os grupos formados e o comportamento observado no alvo. O objectivo é produzir bins que apresentem um padrão consistente, como tendência monotónica ou diferenças claras entre segmentos, pelo que exige ajustes manuais para garantir tamanhos mínimos de amostra e evitar segmentações excessivas.
- **Discretização por árvores de decisão** - Consiste em usar modelos como CART (*Classification and Regression Trees*) para identificar pontos de corte que maximizam a separação entre classes ou reduzem a variância quando o alvo é contínuo (Quinlan, 1996). Conforme destacam Thomas et al. (2017), esse é um dos métodos mais eficientes, pois reduz variabilidade ao mesmo tempo em que preserva o poder discriminativo da variável, e garante que os bins gerados contenham informações relevantes sobre a variável dependente.

2.9.2.6 Selecção de variáveis

A selecção de variáveis consiste em identificar um subconjunto reduzido de atributos que seja ao mesmo tempo necessário e suficiente para representar adequadamente o fenómeno de interesse, usualmente associado a uma tarefa de previsão. Esse processo envolve a escolha de um subconjunto M a partir de um conjunto inicial de N variáveis, com $M < N$, de modo a otimizar uma determinada função de avaliação sobre todos os subconjuntos possíveis de tamanho M (Kira & Rendell, 1992; Narendra & Fukunaga, 1977).

Do ponto de vista predictivo, o objectivo central da selecção de variáveis é reduzir a complexidade do modelo sem comprometer de forma significativa o seu desempenho. Ao eliminar variáveis redundantes ou irrelevantes, busca-se melhorar a eficiência computacional e a capacidade de generalização do algoritmo, utilizando apenas os atributos mais informativos. Dessa forma, o modelo construído com o subconjunto seleccionado tende a aproximar-se de maneira mais eficaz da distribuição da classe original, mantendo elevados níveis de qualidade predictiva (Koller & Sahami, 1996).

Segundo Reddy (2019), Kuhn e Johnson (2013), os principais métodos de selecção de variáveis

podem ser classificados em três categorias:

- **Métodos de filtro:** Os métodos de filtro realizam a selecção de variáveis de forma independente do modelo de aprendizagem, baseando-se em propriedades estatísticas intrínsecas dos dados. Essas abordagens avaliam individualmente a relevância de cada variável em relação à variável alvo, recorrendo a medidas como correlação, informação mútua, estatísticas de dependência ou critérios de variância. Por não envolverem o treinamento do modelo, apresentam baixo custo computacional e boa escalabilidade para conjuntos de dados de grande dimensão (Kira & Rendell, 1992; Kuhn & Johnson, 2013; Reddy, 2019).
- **Métodos wrapper:** Os métodos wrapper avaliam subconjuntos de variáveis de forma iterativa, utilizando directamente o desempenho de um modelo predictivo como critério de selecção. A cada iteração, diferentes combinações de variáveis são testadas por meio do treinamento e validação do modelo, sendo seleccionado o subconjunto que maximiza uma métrica de desempenho previamente definida. Embora tendam a produzir subconjuntos altamente adaptados ao classificador, esses métodos apresentam elevado custo computacional, sobretudo em cenários com grande número de variáveis. Estratégias clássicas incluem selecção progressiva e selecção regressiva (Narendra & Fukunaga, 1977; Koller & Sahami, 1996; Kuhn & Johnson, 2013).
- **Métodos embutidos:** Os métodos embutidos integram o processo de selecção de variáveis directamente na etapa de estimação do modelo. Nessa abordagem, a relevância das variáveis é determinada durante o treinamento, geralmente por meio de penalizações ou critérios internos de optimização. Exemplos clássicos incluem a regressão penalizada do tipo Lasso e modelos baseados em árvores de decisão. Esses métodos apresentam um equilíbrio favorável entre desempenho predictivo, custo computacional e interpretabilidade (Koller & Sahami, 1996; Reddy, 2019).

2.9.2.7 Modelos de machine learning para classificação

Os métodos de aprendizagem supervisionada incluem diversos algoritmos utilizados para resolver problemas de classificação, cada um com diferentes pressupostos e formas de modelar a relação entre as variáveis. Dentre os algoritmos de classificação destacam-se:

2.9.2.7.1 Redes neurais artificiais

As redes neurais artificiais (RNAs) constituem uma classe de modelos estatísticos e computacionais inspirados na organização dos sistemas neurais biológicos. Elas são compostas por unidades elementares, denominadas neurónios artificiais, organizadas em camadas conectadas entre si por pesos ajustáveis. Segundo Hastie et al. (2009), esses modelos procuram

aproximar uma função desconhecida a partir dos dados, sendo particularmente adequados para problemas em que a relação entre os predictores e a variável de interesse apresenta elevada complexidade e não linearidade.

De forma geral, o funcionamento de uma rede neural pode ser compreendido a partir da apresentação de um vector de entrada, ou estímulo, que é propagado sequencialmente pelas camadas da rede até a produção de uma saída. Cada componente do vector de entrada X_i é transmitido às unidades da camada oculta por meio de conexões ponderadas, cujos pesos reflectem a importância relativa de cada variável no processo de decisão do modelo. Esse mecanismo caracteriza as redes neurais como modelos do tipo pós-alimentação *feed forward*, uma vez que a informação flui unidireccionalmente da entrada para a saída.

No contexto de classificação, o objectivo principal é aprender uma função $f_\theta(X)$ que mapeie um vector de predictores $X = (X_1, \dots, X_p)$ para probabilidades associadas a cada classe, minimizando erros de predição (Hastie et al., 2009).

A arquitetura típica utilizada em classificação é o perceptron multicamadas (MLP), composto por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Em uma camada oculta, cada neurónio m calcula inicialmente uma combinação linear das variáveis de entrada, acrescida de um termo de viés, denominada potencial interno ou campo local induzido, dada por:

$$u_m = \alpha_{m0} + \sum_{i=1}^p \alpha_{mi} X_i, \quad m = 1, \dots, M, \quad (2.58)$$

onde α_{mi} representam os pesos associados às entradas X_i e α_{m0} corresponde ao viés do neurónio m . A saída do neurónio oculto é então obtida pela aplicação de uma função de activação não linear $g(\cdot)$ ao potencial interno:

$$Z_m = g(u_m). \quad (2.59)$$

As funções de activação mais utilizadas em MLPs são a função logística (ou sigmoide), definida por:

$$g(u) = \frac{1}{1 + e^{-u}}, \quad (2.60)$$

e a função tangente hiperbólica:

$$g(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}, \quad (2.61)$$

as quais restringem a saída dos neurónios a intervalos finitos, favorecendo a estabilidade numérica do modelo. A introdução dessas não linearidades é fundamental, pois, caso a função de activação seja a identidade, a rede reduz-se a um modelo linear, independentemente do número de camadas.

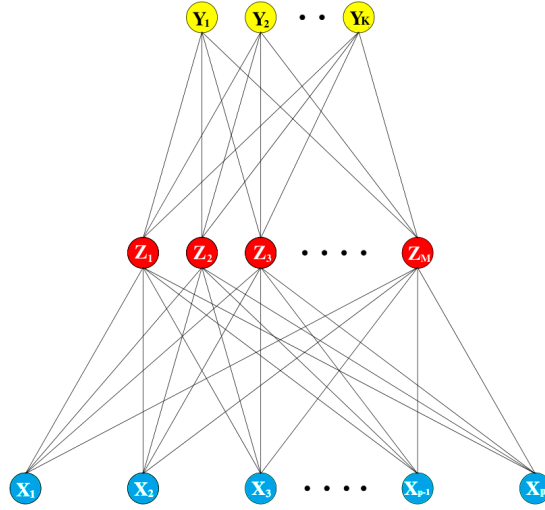


Figura 2.13: Diagrama de uma rede neural com uma camada oculta

Fonte: Hastie et al. (2009)

Na camada de saída, para problemas de classificação multiclasse com K categorias, cada unidade calcula uma pontuação linear da forma:

$$f_k(X) = \beta_{k0} + \sum_{m=1}^M \beta_{km} Z_m, \quad k = 1, \dots, K, \quad (2.62)$$

em que β_{km} são os pesos associados às activações da camada oculta e β_{k0} representa o viés da classe k . Esses escores são posteriormente transformados pela função *softmax*, produzindo probabilidades associadas a cada classe:

$$P(Y = k | X) = \frac{\exp(f_k(X))}{\sum_{j=1}^K \exp(f_j(X))}. \quad (2.63)$$

No caso binário, essa transformação reduz-se à função logística, assegurando valores no intervalo $[0, 1]$ e interpretação probabilística das saídas.

O treinamento da rede consiste em ajustar os pesos das conexões de modo a minimizar uma medida do erro entre as saídas previstas e os valores observados. Em problemas de classificação, utiliza-se comumente a entropia cruzada, definida como:

$$\mathcal{L}(\theta) = - \sum_{i=1}^n \sum_{k=1}^K y_{ik} \log \hat{p}_{ik}, \quad (2.64)$$

onde y_{ik} é uma variável indicadora que assume valor 1 se a observação i pertence à classe k e 0 caso contrário, $\hat{p}_{ik}$ denota a probabilidade estimada pelo modelo e n representa o número total de observações.

A minimização dessa função é realizada por métodos de optimização baseados em gradiente descendente, em que θ denota o vector de todos os parâmetros do modelo, incluindo os pesos e os termos de viés das camadas ocultas e de saída. Nesse procedimento iterativo, os parâmetros do modelo são actualizados na direcção oposta ao gradiente da função de perda, isto é, do vector de derivadas parciais de $\mathcal{L}(\theta)$ em relação aos parâmetros do modelo, avaliado na iteração corrente, segundo a regra:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_{\theta} \mathcal{L}(\theta^{(t)}), \quad (2.65)$$

em que η representa a taxa de aprendizagem e $\nabla_{\theta} \mathcal{L}(\theta^{(t)})$ denota o gradiente da função de perda em relação aos parâmetros do modelo, avaliado na iteração t . Variações estocásticas desse método são amplamente empregadas na prática devido à sua eficiência computacional (Kuhn, 2013).

O cálculo eficiente dos gradientes em redes neurais é viabilizado pelo algoritmo de retropropagação (*backpropagation*), que utiliza a regra da cadeia para propagar o erro da camada de saída para as camadas anteriores. Para a camada de saída, o erro associado à classe k é definido como:

$$\delta_k = \hat{p}_k - y_k, \quad k = 1, \dots, K. \quad (2.66)$$

Em seguida, o erro associado ao neurónio oculto m é dado por:

$$\delta_m = g'(u_m) \sum_{k=1}^K \beta_{km} \delta_k, \quad m = 1, \dots, M, \quad (2.67)$$

onde $g'(u_m)$ denota a derivada da função de activação avaliada no potencial interno. Com isso, o gradiente da função de perda em relação aos pesos da camada oculta assume a forma:

$$\frac{\partial \mathcal{L}}{\partial \alpha_{mi}} = \delta_m X_i, \quad (2.68)$$

considerando-se uma observação individual. No caso do treinamento em lote (*batch*), esses gradientes são somados sobre todas as observações da amostra.

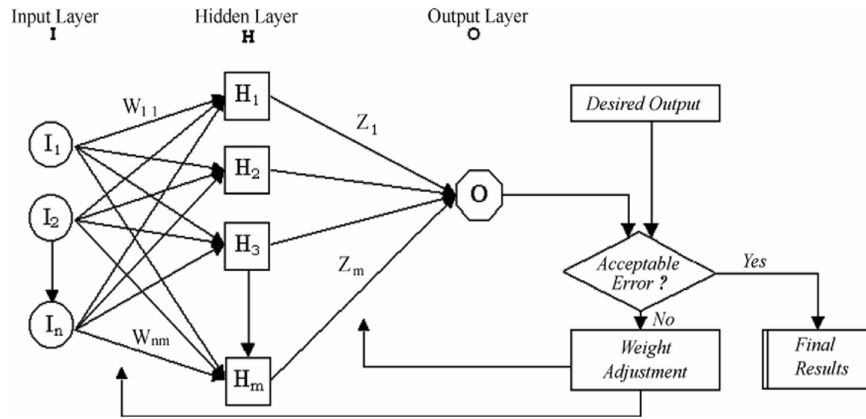


Figura 2.14: Esquema do processo de treinamento de uma rede neural artificial, ilustrando o fluxo feed-forward, o cálculo do erro e o ajuste iterativo dos pesos por retropropagação.

Fonte: Faghihi et al. (2011)

O teorema da aproximação universal estabelece que MLPs com pelo menos uma camada oculta e funções de activação não lineares podem aproximar qualquer função contínua definida em um domínio compacto, desde que disponham de um número suficiente de neurónios (Hastie et al., 2009). Na prática, o número de neurónios ocultos é tratado como um hiper-parâmetro, valores reduzidos podem resultar em subajuste, enquanto valores excessivamente elevados aumentam a complexidade computacional e a dificuldade de optimização. Além disso, a taxa de aprendizagem, a inicialização dos pesos e o número máximo de iterações influenciam directamente a convergência e o desempenho do modelo.

Dessa forma, as RNAs constituem uma ferramenta poderosa e flexível para classificação supervisionada, combinando unidades latentes, funções de activação não lineares e treinamento baseado em gradiente para capturar padrões complexos em conjuntos de dados, apresentando

elevada capacidade discriminativa em contextos caracterizados por relações não lineares entre variáveis (Hastie et al., 2009; Bishop, 2006; Kuhn, 2013).

2.9.2.7.2 Árvore de decisão

As árvores de decisão constituem uma classe de métodos não paramétricos amplamente utilizados em problemas de classificação supervisionada, particularmente adequados a situações em que a relação entre as variáveis predictoras e a resposta é complexa e não linear. Esses algoritmos baseiam-se na partição recursiva do espaço dos predictores em regiões disjuntas, dentro das quais a resposta apresenta comportamento relativamente homogêneo (Hastie et al., 2009; James et al., 2014).

Formalmente, o objectivo de uma árvore de classificação é aprender uma função:

$$f(X) : \mathbb{R}^p \rightarrow \{1, \dots, K\},$$

que associe um vector de predictores $X = (X_1, \dots, X_p)$ a uma das K classes possíveis da variável resposta. Essa associação é realizada por meio de um conjunto hierárquico de regras do tipo *Se-Então*, representadas graficamente por uma estrutura em forma de árvore.

A construção de uma árvore de decisão baseia-se em duas etapas principais: (i) a partição do espaço dos predictores em M regiões disjuntas $R_1, \dots, R_M$; e (ii) a atribuição de uma classe predicta a cada região. No caso da classificação, a classe associada a uma região corresponde, usualmente, à classe mais frequente entre as observações de treinamento alocadas naquela região (Hastie et al., 2009). A partição do espaço dos predictores é realizada por meio do procedimento de divisão binária recursiva. Inicialmente, todas as observações pertencem a uma única região. Em cada passo, busca-se uma variável predictor X_j e um ponto de corte s tais que a divisão do conjunto de dados em duas regiões:

$$R_1(j, s) = \{X : X_j < s\}, \quad R_2(j, s) = \{X : X_j \geq s\},$$

maximize a homogeneidade das classes resultantes.

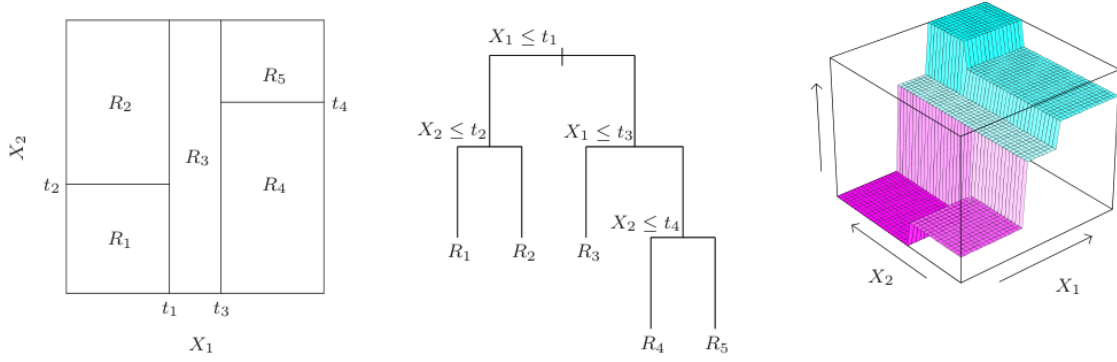


Figura 2.15: Exemplo de particionamento recursivo, árvore correspondente e superfície de predição.

Fonte: Adaptado de Hastie et al. (2009)

Cada região R_m corresponde a um nó terminal (ou folha) da árvore, o qual agrupa um subconjunto das observações de treinamento. No contexto de classificação, a homogeneidade das classes em um nó é quantificada por funções de impureza do nó. Uma das medidas mais simples é a taxa de erro de classificação, definida por:

$$E_m = 1 - \max_k \hat{p}_{mk}, \quad (2.69)$$

Onde $\hat{p}_{mk}$ é a proporção de observações da classe k no nó terminal associado à região R_m .

Outra medida amplamente utilizada é o índice de Gini, dado por:

$$G_m = \sum_{k=1}^K \hat{p}_{mk}(1 - \hat{p}_{mk}), \quad (2.70)$$

o qual pode ser interpretado como uma medida da variabilidade das classes dentro de um nó. Valores baixos de G_m indicam nós mais puros.

Alternativamente, pode-se empregar a entropia cruzada (ou *cross-entropy*), definida por:

$$D_m = - \sum_{k=1}^K \hat{p}_{mk} \log(\hat{p}_{mk}), \quad (2.71)$$

a qual penaliza fortemente nós que contêm observações distribuídas de forma equilibrada entre várias classes (James et al., 2014).

Em cada etapa da divisão binária recursiva, selecciona-se o par (j, s) que minimiza uma medida global de impureza ponderada pelas proporções das observações em cada região

resultante. Esse processo é repetido até que um critério de parada seja atingido, como um número mínimo de observações por nó ou uma profundidade máxima da árvore.

Segundo Hastie et al., (2008) e James et al., (2014), a árvore obtida pelo procedimento de divisão binária recursiva tende a ajustar-se excessivamente aos dados de treinamento, apresentando elevado risco de sobreajuste e, conseqüentemente, pior desempenho predictivo em novas observações. Para mitigar esse problema, emprega-se o procedimento de poda (*pruning*), que consiste em reduzir o tamanho de uma árvore inicialmente completa por meio da remoção controlada de ramos, entendidos como subestruturas da árvore compostas por um nó interno e todos os seus descendentes, resultando em modelos mais simples e com melhor capacidade de generalização.

A poda por complexidade de custo introduz um parâmetro de regularização $\lambda \geq 0$ e define, dentre todas as subárvores possíveis da árvore completa, a subárvore T_λ que minimiza a seguinte função custo:

$$C_\lambda(T) = \sum_{m=1}^{|T|} \sum_{i \in R_m} \mathbb{I}(y_i \neq \hat{y}_{R_m}) + \lambda|T|. \quad (2.72)$$

Nessa expressão, $|T|$ denota o número de nós terminais (folhas) da árvore T , R_m representa a região do espaço dos predictores associada ao m -ésimo nó terminal, y_i é a classe observada da i -ésima observação de treinamento, $\hat{y}_{R_m}$ corresponde à classe predicta para todas as observações pertencentes à região R_m , usualmente definida como a classe mais frequente nessa região, e $\mathbb{I}(\cdot)$ é a função indicadora, que assume valor 1 quando a condição $y_i \neq \hat{y}_{R_m}$ é satisfeita e 0 caso contrário. O termo de penalização $\lambda|T|$ controla o compromisso entre a qualidade de ajuste do modelo e a complexidade da árvore.

À medida que λ aumenta, árvores menores e mais simples são favorecidas, resultando em modelos mais interpretáveis e com melhor capacidade de generalização (Hastie et al., 2009; Kuhn & Johnson, 2013).

Árvores de decisão apresentam diversas vantagens, como interpretabilidade, capacidade de lidar com predictores contínuos e categóricos, robustez à presença de valores ausentes e ausência de pressupostos paramétricos sobre a relação entre as variáveis. No entanto, são modelos instáveis, pois pequenas variações nos dados de treinamento podem resultar em estruturas de árvore substancialmente diferentes (Bishop, 2006).

2.9.2.7.3 Regressão logística

A regressão logística é um modelo estatístico amplamente utilizado em problemas de classificação supervisionada, cujo objectivo é modelar e prever a probabilidade de uma ob-

servação pertencer a uma determinada classe da variável resposta. Apesar de ser denominada regressão, sua aplicação principal está associada a problemas de decisão e discriminação entre classes, especialmente em contextos nos quais a variável resposta é categórica (Hastie et al., 2009; Kuhn & Johnson, 2013).

Considerando uma variável resposta dicotômica, usualmente codificada como $Y \in \{0, 1\}$, assume-se que Y segue uma distribuição binomial condicionada ao vector de predictores $X = (X_1, \dots, X_p)$. Nesse contexto, define-se:

$$p(X) = \Pr(Y = 1 \mid X),$$

como a probabilidade condicional de ocorrência do evento de interesse dado o conjunto de variáveis explicativas.

A principal motivação da regressão logística reside na necessidade de garantir que as probabilidades estimadas pertençam ao intervalo $[0, 1]$, ao mesmo tempo em que se estabelece uma relação sistemática entre a resposta e os predictores. Para isso, o modelo utiliza uma função de ligação não linear, permitindo que combinações lineares das variáveis explicativas sejam mapeadas adequadamente para o espaço de probabilidades.

Para garantir que as probabilidades estimadas pertençam ao intervalo $[0, 1]$ e estejam relacionadas de forma sistemática com os predictores, utiliza-se a função logística, definida por:

$$\Pr(Y = 1 \mid X) = p(X) = \frac{\exp(\beta_0 + \sum_{j=1}^p \beta_j X_j)}{1 + \exp(\beta_0 + \sum_{j=1}^p \beta_j X_j)}. \quad (2.73)$$

Nessa expressão, β_0 representa o intercepto do modelo e β_j são os coeficientes associados aos predictores. A função logística estabelece uma relação não linear entre $p(X)$ e os predictores, embora o modelo seja linear nos parâmetros.

Uma representação equivalente e particularmente útil do modelo é obtida por meio da transformação *logit*, definida como o logaritmo da chance do evento ocorrer:

$$\log \left(\frac{p(X)}{1 - p(X)} \right) = \beta_0 + \sum_{j=1}^p \beta_j X_j. \quad (2.74)$$

Essa forma evidencia que o logaritmo da chance é uma função linear dos predictores, o que

facilita a interpretação dos coeficientes em contextos inferenciais. Em aplicações predictivas, entretanto, o foco recai sobre a capacidade discriminativa do modelo, não sendo estritamente necessários os pressupostos clássicos da inferência estatística (James et al., 2014; Izbicki & Santos, 2018).

O ajuste do modelo de regressão logística é realizado pelo método da máxima verossimilhança, que estima valores para o vector de parâmetros, β . Essas estimativas corresponderão àquelas que resultam em uma probabilidade predicta para cada observação do conjunto de treinamento, $p(x)$, mais próxima da verdadeira classe a qual a observação pertence (James et al., 2014). Considerando um conjunto de treinamento com n observações independentes, a função de verossimilhança é dada por:

$$L(\beta) = \prod_{i=1}^n p(X_i)^{y_i} [1 - p(X_i)]^{1-y_i}, \quad (2.75)$$

em que y_i denota a classe observada da i -ésima observação. As estimativas do vector de parâmetros β são obtidas maximizando-se essa função.

A regra de decisão do classificador logístico é definida a partir da escolha de um ponto de corte para $p(X)$. Um valor usual é $p(X) = 0,5$, de modo que observações com $p(X) > 0,5$ são classificadas como pertencentes à classe 1, enquanto aquelas com $p(X) < 0,5$ são atribuídas à classe 0. Esse ponto de corte define uma fronteira de decisão linear no espaço dos predictores.

Penalização

Em cenários com grande número de predictores ou presença de multicolinearidade, o modelo de regressão logística pode apresentar elevada variância e desempenho predictivo reduzido. Para contornar esse problema, empregam-se técnicas de regularização, que introduzem penalizações à função de verossimilhança com o objectivo de controlar a complexidade do modelo e reduzir o risco de sobreajuste pelo que estes levam a não generalização do modelo em dados não vistos (Hastie, et al. 2009).

Entre as técnicas de regularização mais utilizadas destacam-se *Lasso* (*Least Absolute Shrinkage and Selection Operator*). A penalização do tipo Lasso baseia-se na norma L_1 dos coeficientes e conduz a soluções esparsas, permitindo a selecção automática de variáveis e produzindo modelos mais simples e interpretáveis. A função objectivo da regressão logística penalizada via *Lasso* consiste em minimizar a seguinte função:

$$\hat{\beta} = \arg \min_{\beta} \left\{ - \sum_{i=1}^n [y_i \log p(X_i) + (1 - y_i) \log(1 - p(X_i))] + \lambda \sum_{j=1}^p |\beta_j| \right\},$$

o primeiro termo corresponde à perda associada à verossimilhança binomial. O parâmetro de regularização $\lambda \geq 0$ actua como hiperparâmetro, controlando o grau de penalização, onde valores maiores aumentam a restrição sobre os coeficientes, podendo forçar alguns deles a zero.

O problema de optimização associado à penalização do tipo *Lasso* não admite solução analítica fechada, em virtude da não diferenciabilidade da norma L_1 no ponto zero, sendo, portanto, resolvido por métodos numéricos iterativos, como o algoritmo de descida coordenada (Hastie et al. 2009). A escolha do valor óptimo de λ é geralmente realizada por validação cruzada (procedimento *k-fold*), avaliando o desempenho predictivo do modelo em dados não utilizados no ajuste e garantindo boa generalização para novas observações (Hastie et al. 2009; Kuhn e Johnson, 2013; Raschka, 2017).

2.9.2.8 Avaliação de desempenho de classificadores

Para problemas cuja variável de saída é categórica, utilizam-se diferentes medidas para avaliação dos algoritmos, sendo estas destacadas a seguir.

2.9.2.8.1 Matriz de confusão

A matriz de confusão é uma tabela de contingência que compara os rótulos de classe previstos por um classificador com os rótulos verdadeiros, também chamados de verdade terrestre. Nos casos de classificação binária, a matriz de confusão geralmente assume o formato 2 x 2, apresentando um resumo estatístico que relaciona os valores previstos pelo classificador com os valores observados (Guida, 2018):

	Previsto Negativo	Previsto Positivo
Real Negativo	Verdadeiro Negativo (TN)	Falso Positivo (FP)
Real Positivo	Falso Negativo (FN)	Verdadeiro Positivo (TP)

Tabela 2.2: Matriz de confusão em classificação binária.

De acordo com Guida (2018), a medida mais intuitiva baseada na matriz de confusão é a acurácia, que indica a proporção de classificações correctas em um conjunto de dados, e é calculada da seguinte forma:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (2.76)$$

A precisão indica a proporção de previsões positivas correctas entre todas as previsões positivas.

$$\text{Precisão} = \frac{TP}{TP + FP}. \quad (2.77)$$

A revocação (*recall* ou *sensibilidade*) indica a proporção de verdadeiros positivos identificados correctamente entre todos os casos positivos reais.

$$\text{Sensibilidade} = \frac{TP}{TP + FN}. \quad (2.78)$$

A Especificidade mede a proporção de negativos reais que foram correctamente identificados pelo classificador.

$$\text{Especificidade} = \frac{TN}{TN + FP}. \quad (2.79)$$

O F1-Score representa a média harmónica entre a precisão e a sensibilidade, fornecendo uma medida equilibrada do desempenho do modelo, especialmente útil em situações de desbalanceamento entre as classes.

$$\text{F1-Score} = 2 \times \frac{\text{Precisão} \cdot \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}}. \quad (2.80)$$

2.9.2.8.2 Curva ROC

A curva ROC (*Receiver Operating Characteristic*) é uma ferramenta amplamente utilizada para avaliar o desempenho de classificadores binários (dicotómicos), representando graficamente a relação entre a taxa de verdadeiros positivos (sensibilidade) e a taxa de falsos positivos (1-Especificidade) para diferentes fronteiras de decisão (Fawcett, 2006). A curva fornece uma visão completa do comportamento do classificador, permitindo analisar como a mudança da fronteira afecta a capacidade de discriminação entre as classes.

A medida associada a ROC mais utilizada é a área sob a curva (AUC), que quantifica a habilidade do modelo em separar correctamente classes positivas e negativas. Valores de AUC próximos de 1 indicam excelente poder discriminativo, enquanto valores próximos de 0.5 sugerem desempenho equivalente a um classificador aleatório.

Fawcett (2006), destaca que essa medida é particularmente útil em cenários de classes desbalanceadas, pois avalia a discriminação do modelo ao longo de todas as fronteiras possíveis.

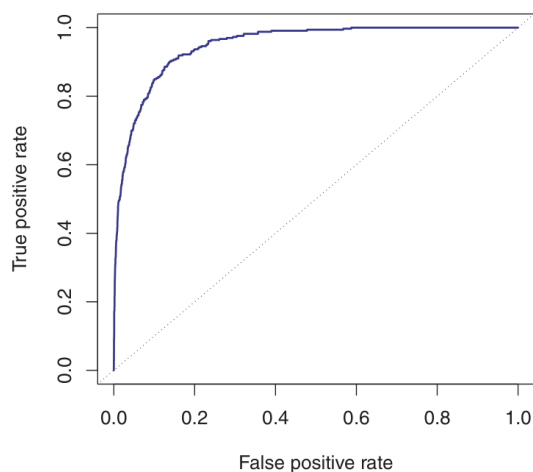


Figura 2.16: Curva ROC

Fonte: Hastie et al. (2009).

As ferramentas de avaliação de desempenho de classificadores, destacadas anteriormente, são abordagens usadas no contexto do modelo individual. Porém, num cenário onde várias abordagens de algoritmos de machine learning aplicados a um mesmo conjunto de observações são desenvolvidos, tendo estes um desempenho muito semelhante ou não, faz-se necessário uma comparação estatística entre estes, a fim de se auxiliar na identificação dos algoritmos que estatisticamente apresentam desempenho superior. Existem várias técnicas com essa finalidade, e neste trabalho são destacados o teste de DeLong e de McNemar.

2.9.2.8.3 Teste de DeLong

O teste de DeLong é um método estatístico não paramétrico utilizado para comparar as áreas sob as curvas ROC (AUC-ROC) entre dois classificadores aplicados ao mesmo conjunto de observações. Proposto por DeLong et al. (1988), como uma alternativa eficiente para avaliar se a diferença entre AUCs é estatisticamente significativa quando as curvas são correlacionadas, isto é, obtidas a partir dos mesmos conjunto de dados.

As hipóteses são:

$$H_0 : AUC_1 = AUC_2 \quad \text{e} \quad H_1 : AUC_1 \neq AUC_2 . \quad (2.81)$$

A estatística do teste é baseada na diferença entre as AUCs estimadas, padronizada pelo

erro padrão dessa diferença:

$$Z = \frac{\hat{AUC}_1 - \hat{AUC}_2}{SE(\hat{AUC}_1 - \hat{AUC}_2)}, \quad (2.82)$$

onde o erro padrão é obtido por:

$$SE(\hat{AUC}_1 - \hat{AUC}_2) = \sqrt{Var(\hat{AUC}_1) + Var(\hat{AUC}_2) - 2Cov(\hat{AUC}_1, \hat{AUC}_2)}, \quad (2.83)$$

A estatística Z segue aproximadamente uma distribuição normal padrão ($Z \sim N(0, 1)$). Assim, rejeita-se a hipótese nula H_0 quando o p-valor associado for menor que o nível de significância α (Robin et al., 2011). Cada AUC é estimada de forma não paramétrica como a probabilidade de o classificador atribuir valor maior a uma amostra positiva do que a uma negativa:

$$\hat{AUC} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \phi(X_i, Y_j), \quad (2.84)$$

em que X_i representa o valor do classificador para a i -ésima observação positiva e Y_j para a j -ésima observação negativa, sendo:

$$\phi(x, y) = \begin{cases} 1, & \text{se } x > y \\ 0.5, & \text{se } x = y \\ 0, & \text{se } x < y \end{cases}. \quad (2.85)$$

Essa formulação permite interpretar a AUC como uma medida de separabilidade entre classes, sendo um indicador robusto da capacidade discriminativa de um classificador (DeLong et al., 1988).

2.9.2.8.4 Teste de McNemar

O teste de McNemar é um método estatístico não paramétrico utilizado para comparar o desempenho de dois classificadores binários aplicados ao mesmo conjunto de observações.

Este método é apropriado quando as previsões de ambos os modelos são dependentes ou pareadas, sendo útil, por exemplo, para verificar se há diferença significativa nas proporções de acertos entre os dois métodos (McNemar, 1947).

O teste se baseia em uma tabela de contingência 2×2 , na qual apenas as observações discordantes entre os classificadores são relevantes para o cálculo do teste.

Tabela 2.3: Tabela de contingência para o Teste de McNemar

	Resultado B = 1	Resultado B = 0	Total
Resultado A = 1	n_{11}	n_{10}	$n_{1\cdot}$
Resultado A = 0	n_{01}	n_{00}	$n_{0\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 0}$	n

A hipótese nula do teste é que os dois classificadores possuem o mesmo desempenho, ou seja, as proporções de discordâncias são iguais, isto é:

$$H_0 : p_{01} = p_{10} \quad \text{e} \quad H_1 : p_{01} \neq p_{10}. \quad (2.86)$$

O teste estatístico de McNemar clássico é calculado da seguinte forma:

$$\chi^2 = \frac{(n_{01} - n_{10})^2}{n_{01} + n_{10}}, \quad (2.87)$$

e segue, aproximadamente, uma distribuição qui-quadrado com um grau de liberdade (Agresti, 2007). Quando o número de discordâncias $n_{01} + n_{10}$ é pequeno (geralmente menor que 25), recomenda-se aplicar a correção de continuidade de Yates, obtendo-se o teste estatístico ajustado:

$$\chi^2 = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}}, \quad (2.88)$$

O critério de decisão é baseado na comparação entre o valor calculado e o valor crítico da distribuição qui-quadrado com 1 grau de liberdade, ao nível de significância α (geralmente 5%). Assim:

$$\text{Rejeita-se } H_0 \text{ se } \chi^2_{\text{calculado}} > \chi^2_{\text{crítico}}(1; \alpha), \quad (2.89)$$

ou, de forma equivalente, rejeita-se H_0 quando o p-valor associado for menor que α . Caso H_0 seja rejeitada, conclui-se que existe diferença estatisticamente significativa entre os classificadores, ou seja, um dos modelos apresenta desempenho superior. Se H_0 não for rejeitada, entende-se que não há evidências suficientes para afirmar que as proporções de acerto diferem.

No contexto da avaliação de desempenho de um único modelo, o teste de McNemar pode ser aplicado entre as classificações observadas e estimadas. Nesse caso, o teste avalia se há uma diferença sistemática entre as proporções de classificações incorrectas para cada classe, ou seja, se o modelo apresenta viés em relação a uma das categorias da variável resposta. Neste contexto, se não rejeitamos H_0 , o modelo não apresenta viés sistemático entre as classes, indicando que as proporções de erros tipo I e tipo II são semelhantes. Se rejeitamos H_0 , o modelo tende a classificar uma das classes com maior taxa de erro, o que sugere desequilíbrio no desempenho (por exemplo, um algoritmo que classifica melhor os casos positivos do que os negativos).

2.9.2.9 Implementação de melhoramento do modelo

Uma etapa crítica no fluxo de trabalho de machine learning sucede o treinamento inicial), no qual frequentemente utilizam-se hiperparâmetros padrão ou ajustes manuais preliminares. Para que o modelo atinja seu potencial máximo, é necessário realizar o ajuste fino (*fine tuning*), visando maximizar o desempenho predictivo e assegurar a generalização. Conforme pontuado por Géron (2019), a otimização de modelos é fundamental não apenas para a precisão, mas também para a eficiência computacional. No contexto financeiro, López de Prado (2018) enfatiza que, em finanças, a otimização deve ser acompanhada de validações rigorosas para evitar resultados ilusórios. Algumas das abordagens de melhoramento de algoritmos são descritas a seguir.

2.9.2.9.1 Hiperparametrização dos modelos de machine learning

Diferente dos parâmetros dos modelos, que são estimados internamente a partir dos dados durante o treinamento (como pesos em regressão logística), os hiperparâmetros são configurações externas definidas antes do início do processo de aprendizagem (Ayub et al., 2022). Eles actuam como botões de controle que determinam a complexidade do modelo e a velocidade de convergência. A literatura destaca três abordagens principais para a busca da configuração dos hiperparâmetros, também conhecido como *fine tuning*:

- **Pesquisa em grade (*grid search*):** Caracteriza-se como uma busca exaustiva e determinística dentro de um espaço de busca previamente delimitado. O algoritmo testa todas as combinações possíveis entre os valores sugeridos (Klatzer e Pock, 2015). Embora garanta a identificação da melhor combinação dentro da grade estipulada, sua

principal desvantagem é a maldição da dimensionalidade, tornando-se computacionalmente inviável à medida que o número de hiperparâmetros aumenta.

- **Pesquisa aleatória (*random search*):** Em contraste com a exaustão da grade, esta técnica selecciona combinações aleatórias dentro de distribuições estatísticas definidas para cada hiperparâmetro. Bergstra e Bengio (2012), demonstraram que a pesquisa aleatória é frequentemente mais eficiente que a pesquisa em grade, pois consegue explorar o espaço de busca de forma mais ampla com um custo computacional significativamente menor, identificando ótimos globais que a grade poderia ignorar.

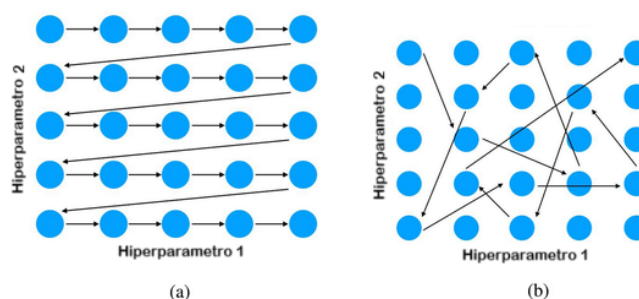


Figura 2.17: (a) Grid search; (b) Random Search.

Para que estas técnicas de busca sejam aplicadas a séries temporais, a literatura recomenda que o cálculo da medida de desempenho de cada combinação utilize a Validação Cruzada K-Fold para Séries Temporais. Diferente do k -fold tradicional, onde os dados são embaralhados, esta adaptação respeita a dependência cronológica, o processo inicia-se com um subconjunto de treino que é ampliado a cada *fold*, enquanto o teste é sempre o bloco de dados imediatamente subsequente (Packt, 2019).

Dessa forma, tanto o *grid search* quanto o *random search* operam sobre as dobras temporais para garantir que a escolha dos melhores hiperparâmetros não sofra de vazamento de dados, assegurando que o modelo nunca utilize informações do futuro para validar previsões do passado (Hyndman & Athanasopoulos, 2021).

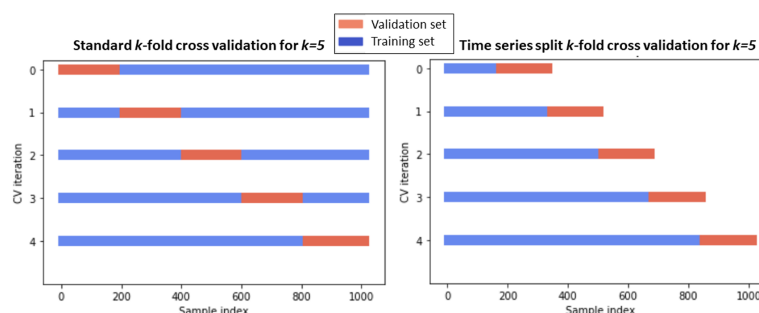


Figura 2.18: Validação cruzada k-fold clássica vs validação cruzada de série temporal.

Fonte: Packt (2019).

Após a optimização dos hiperparâmetros, obtém-se uma configuração que maximiza o desempenho segundo uma medida pré-definida. No entanto, conforme destacado por Géron (2019), o desempenho durante o ajuste não garante, por si só, capacidade de generalização. Em problemas envolvendo séries temporais financeiras, torna-se, portanto, necessário proceder a uma avaliação fora da amostra que respeite a ordem cronológica dos dados (López de Prado, 2018; Hyndman & Athanasopoulos, 2021). Nesse contexto, destacam-se os métodos apresentados nas subsecções seguintes.

2.9.2.9.2 Validação cruzada em séries temporais

A validação cruzada constitui uma técnica estatística amplamente utilizada para estimar o erro de generalização de modelos predictivos, fornecendo uma avaliação mais fiável do seu desempenho em dados não observados (Arlot & Celisse, 2010). Contudo, no contexto de séries temporais, a aplicação directa dos métodos tradicionais baseados em permutação aleatória revela-se inadequada, uma vez que ignora a dependência temporal entre as observações, podendo introduzir *data leakage* e enviesar a estimativa do erro de previsão (Hyndman & Athanasopoulos, 2021).

Nesse contexto, destaca-se a validação cruzada para séries temporais, também designada por *Walk-Forward Validation*, que se baseia na avaliação do modelo em múltiplas origens de previsão ao longo do tempo. Em cada iteração, o modelo é ajustado exclusivamente com dados passados e avaliado em observações futuras, respeitando a estrutura temporal da série. De acordo com Cerqueira et al. (2019), esta abordagem fornece uma estimativa mais robusta do desempenho predictivo, por considerar diferentes períodos da série.

Dentro desta classe de métodos, distinguem-se:

- **Validação cruzada com janela expansiva:** Nesta abordagem, o conjunto de treino cresce progressivamente à medida que novas observações são incorporadas. Em cada

iteração, o modelo é ajustado utilizando todas as observações disponíveis até um determinado instante temporal e testado no período subsequente. Segundo Hyndman e Athanasopoulos (2021), este método é particularmente adequado quando se assume que a informação histórica permanece relevante para a previsão futura, permitindo um processo de aprendizagem cumulativa ao longo do tempo.

- **Validação cruzada com janela deslizante:** Nesta variante, o conjunto de treino possui tamanho fixo e desloca-se ao longo do tempo, descartando observações mais antigas à medida que novas observações são adicionadas. Embora seja útil em contextos de não estacionariedade, este método pode resultar na perda de informação histórica relevante, sendo menos apropriado quando se pretende explorar toda a amostra disponível (Cerqueira et al., 2019).

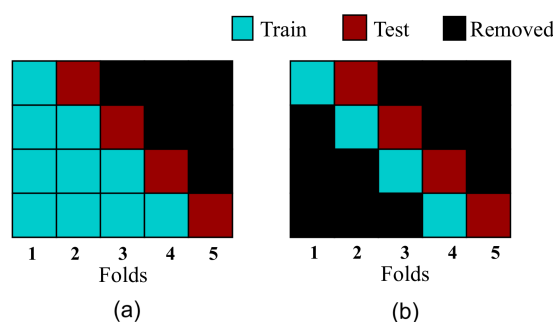


Figura 2.19: Validação Cruzada com: (a) Janela Expandida; (b) Janela Deslizante.

Fonte: Cerqueira et al. (2019).

2.9.2.9.3 Bootstrap em séries temporais

O bootstrap é uma técnica estatística de reamostragem amplamente utilizada para aproximar a distribuição amostral de estimadores, permitindo a estimação de medidas de incerteza, como viés, variância e intervalos de confiança, sem recorrer a suposições paramétricas fortes. Introduzido por Efron (1979), o método baseia-se na geração de múltiplas amostras obtidas por reamostragem com reposição a partir da amostra original, tratando-a como uma aproximação da população subjacente.

No contexto clássico, o bootstrap assume que as observações são independentes e identicamente distribuídas (i.i.d.). Contudo, esta suposição é violada em séries temporais, uma vez que observações consecutivas tendem a apresentar dependência temporal. Como discutem Davison e Hinkley (1997) e Lahiri (2003), a aplicação directa do bootstrap clássico a dados temporais pode conduzir a estimativas enviesadas da variabilidade dos estimadores, por destruir a estrutura de dependência serial dos dados.

Em resposta a essa limitação, a literatura propôs extensões do bootstrap especificamente adaptadas a séries temporais, as quais procuram preservar, ao menos parcialmente, a dependência temporal entre observações consecutivas. Uma das abordagens fundamentais é o bootstrap por blocos, introduzido por Künsch (1989), que consiste na divisão da série em blocos consecutivos de observações, reamostrados com reposição para formar novas séries bootstrap, preservando assim a dependência temporal local.

Dentro da família do bootstrap por blocos, destacam-se algumas variantes:

- **Bootstrap por blocos móveis:** Esta extensão define blocos sobrepostos, extraídos de uma janela deslizante ao longo da série. A sobreposição aumenta o número de blocos disponíveis para reamostragem, proporcionando estimativas mais estáveis. Conforme Lahiri (2003), o método apresenta melhores propriedades assintóticas em comparação com o bootstrap por blocos não sobrepostos. Segundo Gilda et al. (2024), essa técnica é especialmente útil na análise de séries temporais financeiras, onde a captura das dependências de curto prazo é crucial para previsões precisas.
- **Bootstrap estacionário:** Proposto por Politis e Romano (1994), o bootstrap estacionário utiliza blocos de comprimento aleatório, geralmente gerados a partir de uma distribuição geométrica. Essa estratégia elimina a necessidade de fixar um tamanho de bloco específico e garante que as séries bootstrap geradas sejam estacionárias, sendo particularmente útil quando se assume estacionariedade fraca da série temporal.

2.10 Peso da evidência

O método Peso da Evidência ou *Weight of Evidence* (WOE), é uma técnica estatística amplamente utilizada em modelagem supervisionada, especialmente em contextos de análise de risco e previsão de eventos binários, como crédito e fraude. O conceito foi originalmente desenvolvido no contexto da teoria da informação e posteriormente adaptado para modelagem estatística (Good, 1950; Anderson, 2007; Siddiqi, 2012).

O WOE tem como principal objectivo quantificar a força e a direcção da associação entre uma variável explicativa e uma variável alvo binária, convertendo categorias (ou intervalos numéricos) em valores numéricos contínuos. Essa transformação é particularmente útil para modelos lineares, como regressão logística, pois substitui as variáveis originais por medidas que reflectem a evidência estatística de separação entre as classes.

A ideia central desta técnica é comparar a proporção de ocorrências de cada classe (por exemplo, sucesso e fracasso) dentro de cada intervalo (bin) de uma variável explicativa.

A fórmula do WOE para o *bin* i é dada por:

$$\text{WOE}_i = \ln \left(\frac{\text{Distribuição dos eventos positivos no bin } i}{\text{Distribuição dos eventos negativos no bin } i} \right), \quad (2.90)$$

ou, de forma expandida:

$$\text{WOE}_i = \ln \left(\frac{\frac{B_i}{B_T}}{\frac{M_i}{M_T}} \right), \quad (2.91)$$

onde:

- B_i : número de observações positivas no bin i ;
- M_i : número de observações negativas no bin i ;
- B_T : total de observações positivas no conjunto de dados;
- M_T : total de observações negativas no conjunto de dados;

O valor de WOE_i pode ser interpretado da seguinte forma:

- $\text{WOE}_i > 0$: indica que o bin tem uma maior proporção de eventos positivos (ou sucesso);
- $\text{WOE}_i < 0$: indica uma maior proporção de eventos negativos (ou fracasso);
- $\text{WOE}_i = 0$: indica ausência de poder discriminante entre as classes nesse bin.

Essa transformação é especialmente útil porque padroniza as variáveis explicativas, permitindo comparabilidade, facilita a interpretação dos coeficientes em modelos logísticos, e melhora a estabilidade temporal do modelo, tornando-o mais robusto frente a pequenas mudanças na distribuição dos dados (Siddiqi, 2012).

2.11 MetaTrader5

O MetaTrader 5 foi oficialmente lançado em 1º de junho de 2010 pela MetaQuotes como uma plataforma de negociação multi-mercado projectada para atender tanto ao Forex quanto a

mercados de acções, opções, entre outros. Esta tornou-se uma das plataformas de negociação mais difundidas entre correctoras e participantes do mercado financeiro mundial, devido à sua versatilidade e integração com diversos intermediários. O software disponibiliza ao utilizador ferramentas para visualização gráfica em tempo real, aplicação de indicadores técnicos e execução de operações de compra e venda (MetaQuotes, 2025a).

A plataforma distingue-se também pela incorporação do sistema Expert Advisors (EAs), que possibilita a automatização de estratégias de negociação. Essas estratégias são estruturadas por meio de algoritmos de decisão. No desenvolvimento desses algoritmos no MetaTrader 5 é utilizado a linguagem MetaQuotes 5 (MQL5).

Paralelamente ao lançamento da plataforma, surgiu uma comunidade especializada em programação MQL5. Esta comunidade reúne desenvolvedores e traders, oferecendo um espaço para comercialização e aluguel de algoritmos, bem como uma colecção abrangente de artigos e materiais técnicos dedicados à linguagem.

2.12 Trabalhos relacionados

A literatura recente evidencia um crescente interesse na aplicação de técnicas de machine learning para a previsão direccional de retornos no mercado Forex, bem como na conversão dessas previsões em estratégias operacionais.

Baasher e Fakhr (2010), realizaram um estudo para prever a direcção do preço dos pares USD/JPY, USD/EGP e EUR/EGP utilizando dados diários num período de 1852 dias, de abril de 2003 a agosto de 2010. Avaliaram três modelos: rede neural de função de base radial (RBF), rede neural perceptron de multiplas camadas (MLP) e máquina de vectores de suporte (SVM). Os modelos foram treinados utilizando 11 indicadores técnicos como variáveis de entrada. Entre eles, o modelo MLP alcançou a maior precisão, atingindo 79,2%.

Recentemente, Enkhbayar e Ślepaczuk (2025), desenvolveram um sistema abrangente para geração de sinais de negociação (trading) baseado em machine learning. O estudo comparou modelos como regressão ridge, KNN, floresta aleatória, XGBoost, redes neurais artificiais, LSTM e GRU em dados diários e dentro do dia. Foi utilizada optimização *walk forward* e análise por janelas móveis. Os resultados mostraram que modelos de *ensemble*, especialmente XGBoost e a floresta aleatória, obtiveram melhor desempenho predictivo e económico. As estratégias baseadas nesses modelos apresentaram índices de Sharpe superiores as referências técnicas tradicionais, indicando maior eficiência na gestão risco retorno.

Chihab et al. (2019), propuseram um sistema para negociar no mercado Forex de ao longo da semana, baseado numa combinação de alguns dos indicadores técnicos mais comuns. O

sistema foi implementado combinando um modelo de regressão probit e um algoritmo de floresta aleatória, alcançando um ROI de 78% ao longo de 17 semanas no par de moedas GBP/USD.

López-Herrera et al. (2025), investigaram a previsão direccional em oito pares cambiais utilizando regressão logística, árvores de decisão, redes neurais e métodos de aumento de gradiente. A avaliação foi conduzida com validação temporal e análise de sensibilidade a diferentes horizontes de previsão. Os autores observaram que o XGBoost apresentou a maior acurácia média, enquanto a regressão logística regularizada demonstrou maior robustez em períodos de elevada volatilidade. Em termos económicos, os modelos produziram retornos positivos quando integrados em estratégias baseadas em limiares probabilísticos, confirmando a utilidade prática da classificação.

Por sua vez, Jubert de Almeida et al. (2018), propuseram um sistema híbrido que combina uma máquina de vectores de suporte com algoritmos genéticos (GA). Nesse sistema, os algoritmos genéticos são utilizados para otimizar regras de negociação baseadas em diferentes indicadores técnicos. A SVM, por sua vez, é responsável por classificar previamente o tipo de mercado no momento actual. Com base nessa classificação, o sistema selecciona um de três algoritmos genéticos treinados de forma distinta, cada um adaptado a um regime específico de mercado. O melhor modelo proposto pelos autores alcançou um ROI de aproximadamente 83% num período ligeiramente superior a um ano.

Thu e Xuan (2018), investigaram a aplicação de máquina de vectores de suporte (SVM) na previsão da direcção dos movimentos do mercado cambial, formulando o problema como uma tarefa de classificação binária entre movimentos ascendentes e descendentes. O modelo foi utilizado para negociar o par EUR/USD durante o ano de 2016, sendo implementada uma estratégia simples baseada nos sinais gerados pela SVM. Adicionalmente, os autores integraram o classificador num robô de negociação para a plataforma MetaTrader 4, permitindo a execução automatizada das operações. Os resultados demonstraram que o sistema baseado em SVM apresentou desempenho superior em relação a estratégias sem o SVM, alcançando um retorno sobre o investimento de aproximadamente 33,8%, evidenciando o potencial da abordagem para apoiar decisões de negociação no mercado Forex.

Guyard e Deriaz (2024) testaram vários algoritmos de machine learning para prever a direcção do par EUR/USD utilizando dados diários. Entre todos os modelos avaliados, o *histogram based gradient boosting* alcançou a maior precisão (58%) e lucro (29.5%).

Capítulo 3

MATERIAL E METODOS

O material usado e as técnicas aplicadas para o alcance dos objectivos e resultados deste trabalho são expostos neste capítulo.

3.1 Material

Para o desenvolvimento deste estudo, foram seleccionadas diferentes ferramentas de apoio. Utilizou-se a linguagem R versão 4.4.1, em conjunto com as suas bibliotecas e a plataforma RStudio. A escolha do R justifica-se pela sua simplicidade, pela vasta gama de bibliotecas disponíveis para machine learning e manipulação de dados, bem como pelo facto de se tratar de uma linguagem de código aberto. As análises estatísticas neste estudo foram conduzidas com um nível de significância de 5%, utilizando o valor p como critério para a tomada de decisão nos testes de hipóteses.

Adicionalmente, recorreu-se ao TeXstudio para a edição do relatório em $\text{\LaTeX}$, ao Microsoft Excel 2016 para a pré-elaboração das tabelas.

O estudo será aplicado nos pares do mercado de Forex, especificamente para os seguintes pares de moedas:

- EUR/USD (Euro / Dólar Americano)
- USD/JPY (Dólar Americano / Iene Japonês)
- GBP/USD (Libra Esterlina / Dólar Americano)

Os dados utilizados estão no formato CSV (*Comma separated values*) e foram obtidos directamente do MetaTrader 5, uma das principais plataformas de negociação mais utilizadas para operações no mercado Forex. O dataframe inclui dados diários brutos dos preços de CFDs sobre pares de moedas, com informações sobre os preços de abertura, máxima, mínima, fechamento, o volume e o volume de ticks, de onde novas variáveis foram derivadas (Tabela 3.1). Sendo o alvo ou variável dependente RA_t , a classificação dos retornos do par de moeda no dia seguinte. O retorno R_t é definido pela seguinte expressão:

$$R_t = \frac{C_t - C_{t-1}}{C_{t-1}}, \quad (3.1)$$

a partir do qual a variável resposta RA_t (Tabela 3.1) é definida como:

$$RA_t = \begin{cases} 1 & \text{(Compra), se } R_{t+1} \geq 0, \\ 0 & \text{(Venda), caso contrário.} \end{cases} \quad (3.2)$$

Onde:

- R_{t+1} é o retorno do par de moedas no período $t + 1$,
- C_t é o preço de fechamento do par de moedas no final do período t ,
- C_{t-1} é o preço de fechamento do par de moedas no final do período anterior $t - 1$.

Os dados utilizados correspondem aos períodos de 09/09/2010 a 31/01/2025. Para cada par de moedas, as variáveis apresentadas na Tabela 3.1 foram calculadas em diferentes janelas temporais de 5, 10, 15, 20 e 52 períodos, pelo que totalizam 209 variáveis.

Nomenclatura das variáveis

As 209 variáveis explicativas utilizadas neste estudo seguem uma convenção de nomenclatura padronizada, definida de forma a identificar explicitamente o tipo de variável, o horizonte temporal considerado e, quando aplicável, o activo financeiro de referência. A notação adotada é descrita a seguir:

- **Nome da variável + período:** Representa variáveis calculadas com base em uma janela temporal fixa. Por exemplo, VA15 denota a volatilidade anualizada estimada a partir de uma janela de 15 períodos. Um segundo tipo adiciona ao final o nome do

activo específico pelo qual a variável é calculada; por exemplo, VA15EURUSD, que denota a volatilidade anualizada do EUR/USD estimada a partir de uma janela de 15 períodos.

- **Nome da variável + período + activo secundário:** Utilizada para variáveis que envolvem a relação entre o retorno do par de moeda e de outro activo. Por exemplo, CEA10ROIL representa a correlação entre o par cambial analisado e do petróleo (ROIL), calculada em uma janela de 10 períodos.
- **Nome da variável + período primário + período da variável secundária:** Empregada quando a variável é construída a partir de retornos agregados em um horizonte distinto da janela de estimação. Por exemplo, VA5.10 corresponde à volatilidade anualizada calculada em uma janela de 5 períodos, com base em retornos de 10 períodos.

Tabela 3.1: Descrição das variáveis

ID	Descrição da Variável	Tipo de Variável
RSP500	Retorno do S&P 500	Quantitativa contínua
VRSP500	Volatilidade Anualizada dos Retornos S&P500	Quantitativa contínua
RDX	Retorno do DXY (Índice do Dólar)	Quantitativa contínua
VDXY	Volatilidade Anualizada do DXY	Quantitativa contínua
RSTOXX	Retorno do Euro Stoxx 50	Quantitativa contínua
VSTOXX	Volatilidade Anualizada do Euro Stoxx 50	Quantitativa contínua
RNK225	Retorno do Nikkei 225	Quantitativa contínua
VNK225	Volatilidade Anualizada do Nikkei 225	Quantitativa contínua
ROIL	Retorno do Petróleo Bruto (WTI)	Quantitativa contínua
VOIL	Volatilidade Anualizada do Petróleo Bruto	Quantitativa contínua
HUSRT	Coefficiente de Hurst	Quantitativa contínua
ASMTR	Assimetria dos Retornos	Quantitativa contínua
CURTOSE	Curtose dos Retornos	Quantitativa contínua
CVR	Coefficiente de Variação dos Retornos	Quantitativa contínua
PER	Probabilidade de Estacionariedade dos Retornos	Quantitativa contínua
LRY	Retornos Defasados	Quantitativa contínua
%B	Percentual de Bollinger (%B)	Quantitativa contínua
IVM	Intervalo Verdadeiro Médio	Quantitativa contínua
VA	Volatilidade Anualizada	Quantitativa contínua
EVRS	Volatilidade Rogers-Satchell	Quantitativa contínua
IFR	Índice de Força Relativa	Quantitativa contínua
OE	Oscilador Estocástico	Quantitativa contínua
TM	Taxa de Mudança dos Activos	Quantitativa contínua
RFA	Razão entre o Preço Máximo e o Preço Mínimo	Quantitativa contínua
RFC	Proporção do Corpo do Candle	Quantitativa contínua
CEA	Correlação entre o Par e outros Activos	Quantitativa contínua
W%	Williams %R	Quantitativa contínua
OMC	Oscilador de Momento de Chande	Quantitativa contínua
IFDC	Fluxo de Dinheiro de Chaikin	Quantitativa contínua
IMD	Movimento Direccional	Quantitativa contínua
RA	Classificação dos Retornos no Dia Seguinte	Binária

3.2 Métodos

As variáveis na tabela 3.1, foram extraídas dos dados brutos, através do processo de extracção de atributos, também conhecida como engenharia de características, processo que transforma dados brutos em variáveis mais informativas, capazes de melhorar o desempenho dos modelos predictivos (Reddy, 2019). Este processo envolveu a utilização de janelas temporais móveis, o que resultou em observações iniciais com valores ausentes. Adicionalmente, a integração da base de dados dos pares de moedas com os índices, que possuíam menor número de observações, contribuiu para a redução do tamanho final do conjunto de dados. Todos os valores ausentes foram posteriormente removidos. Ressalta-se que os conceitos e definições das variáveis encontram-se apresentados na Secção 2.

3.2.1 Análise exploratória de dados

A análise exploratória neste estudo teve como finalidade caracterizar o comportamento dos pares cambiais seleccionados. Neste recorreu-se a representações gráficas, como histogramas, boxplots, gráficos de dispersão e mapas de calor, para caracterizar a distribuição dos dados. Empregaram-se também testes estatísticos em janelas móveis para avaliar algumas propriedades das séries temporais (detalhado em 2.6.2), utilizando recursos de visualização gráfica.

3.2.2 Divisão do conjunto de dados

Após a análise exploratória, os dados foram divididos em dois subconjuntos: treino (70%) e teste (30%), correspondentes aos períodos de 04.02.2013 a 02.02.2022 e de 03.02.2022 a 31.01.2025, respectivamente.

As tabelas 3.2 e 3.3 expressam o número de observações nas bases de treinamento e teste.

Classe/Par	Compra	Venda
USD/JPY	1097 (49,36%)	1125 (50,63%)
GBP/USD	1120 (50,22%)	1110 (49,77%)
EUR/USD	933 (50,78%)	904 (49,21%)

Tabela 3.2: Base de treino,

Classe/Par	Compra	Venda
USD/JPY	403 (44,09%)	511 (55,90%)
GBP/USD	450 (49,28%)	463 (50,71%)
EUR/USD	465 (50,93%)	448 (49,06%)

Tabela 3.3: Base de teste.

3.2.3 Pré-processamento de dados

Esta etapa concentrou-se na base de treinamento, onde todas as variáveis explicativas passaram por procedimentos de pré-processamento com o objectivo de melhorar a qualidade e a interpretabilidade dos dados utilizados nos modelos de previsão.

Aplicou-se a discretização supervisionada por meio de árvores de decisão, a qual agrupou os valores contínuos das variáveis em bins definidos com base na relação estatística com a variável alvo.

Posteriormente, cada bin resultante foi transformado utilizando o método peso da evidência (*WOE*), atribuindo valores numéricos que expressam a força e a direcção da associação entre cada bin e a variável alvo.

3.2.4 Critério de selecção de variáveis

A selecção de variáveis também foi realizada na base de treinamento, com foco na redução da multicolinearidade entre os predictores. Para isso, aplicou-se um método heurístico de remoção iterativa de variáveis altamente correlacionadas, estratégia para manter apenas o conjunto mínimo de predictores não redundantes (Kuhn & Johnson, 2013). A matriz de correlação entre todas as variáveis independentes foi calculada utilizando o coeficiente de Spearman. Foram definidos pontos de corte empíricos de correlação absoluta de $|r_k| < 0,25$ para EUR/USD e GBP/USD, e $|r_k| < 0,40$ para USD/JPY. O procedimento do critério de selecção das variáveis está descrito no Algoritmo 1.

Algorithm 1: Critério de Selecção de Variáveis: Remoção Heurística de predictores Altamente Colineares

Data: Matriz de predictores $X_1, \dots, X_p$

Result: predictores sem multicorrelação acima de k

Calcular a matriz de correlação

Definir ponto de corte k

repeat

 Identificar pares com maior correlação absoluta ($|r| > k$)

 Seja o par (X_i, X_j) multicorrelacionado

 Calcular correlação média de X_i e X_j com os demais

if correlação média de $X_i > X_j$ **then**

 Remover X_i

else

 Remover X_j

end

until Todas as correlações absolutas $< k$

Fonte: Kuhn e Johnson (2013)

3.2.5 Algoritmos de machine learning

A fim de se fazer a classificação da direcção dos retornos futuros para cada par, utilizou-se três algoritmos de machine learning.

Esse processo de treinamento dos algoritmos foi conduzido separadamente para cada um dos três pares de moedas analisados e estruturado em duas etapas principais. Na primeira etapa, os algoritmos foram treinados utilizando configurações de hiperparâmetros definidas de forma empírica, manual ou a partir dos valores padrão da implementação em R, com o objectivo de estabelecer um modelo de referência para o desempenho predictivo. Essa etapa

inicial permitiu obter uma avaliação preliminar da capacidade de classificação dos modelos, antes da aplicação de procedimentos sistemáticos de optimização de hiperparâmetros (3.2.7).

Regressão logística penalizada

Utilizou-se a regressão logística com penalização Lasso. O modelo estima a probabilidade de a variável dependente y_t (directão do retorno no período seguinte) assumir o valor 1 (compra) ou 0 (venda), dado o vector de predictores X_t disponível no momento actual.

A estimação dos parâmetros foi conduzida através da minimização da função de custo, que integra a penalidade ℓ_1 à log-verossimilhança:

$$\min_{\beta_0, \beta} \left\{ -\frac{1}{n-1} \sum_{t=1}^{n-1} [y_t(\beta_0 + x_t^T \beta) - \log(1 + e^{\beta_0 + x_t^T \beta})] + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (3.3)$$

Onde λ é o parâmetro de regularização que controla a intensidade da penalidade.

Árvore de decisão

Também utilizou-se o modelo de árvore de decisão para a tarefa de classificação. O algoritmo segmenta o espaço dos predictores X_t em regiões retangulares disjuntas, atribuindo a cada região a classe majoritária da variável dependente y_t .

O ajuste do modelo foi realizado por meio de divisão binária recursiva, adotando o índice de Gini como critério de impureza para a selecção das partições óptimas.

O índice de Gini para um nó m é definido por:

$$G = \sum_{k=1}^K \hat{p}_{mk}(1 - \hat{p}_{mk}) \quad (3.4)$$

Onde $\hat{p}_{mk}$ representa a proporção de observações da classe k no nó m .

Para evitar o sobreajuste, o crescimento da árvore foi controlado por um critério de penalização da complexidade, que interrompe a criação de novos nós quando o ganho de informação não atinge um limiar mínimo de desempenho, assegurando a parcimónia do modelo (Breiman et al., 1984; Kuhn & Johnson, 2013).

Rede neural artificial

Utilizou-se a rede neural artificial do tipo perceptron multicamadas (MLP), com arquitetura *feed forward* e multiplas camadas ocultas, aplicada à tarefa de classificação binária. Esse tipo de arquitetura é amplamente utilizado em problemas de classificação supervisionada devido à sua capacidade de modelar relações não lineares entre as variáveis explicativas e a variável resposta, mesmo em contextos de elevada complexidade estrutural (Hastie et al., 2009).

Utilizou-se uma rede neural artificial do tipo perceptron multicamadas (MLP). A arquitetura implementada é do tipo *feed forward*, processando o vector X_t através de camadas ocultas para prever a direcção do retorno no dia seguinte.

A estimação dos pesos sinápticos foi conduzida por meio do algoritmo de backpropagation, utilizando a minimização da função de perda de entropia cruzada, definida por:

$$L(\mathbf{w}) = -\frac{1}{n} \sum_{t=1}^n [y_t \log(\hat{y}_t) + (1 - y_t) \log(1 - \hat{y}_t)] \quad (3.5)$$

Onde $\hat{y}_t$ representa a probabilidade estimada pela rede. Utilizou-se a função de activação logística (sigmóide) tanto nas unidades ocultas quanto na camada de saída, garantindo que o sinal resultante seja mapeado no intervalo $[0, 1]$. Este mapeamento permite a interpretação das saídas como probabilidades condicionais de alta (compra) ou baixa (venda) para cada par de moedas. O processo de aprendizagem foi realizado de forma iterativa via gradiente descendente.

3.2.6 Medidas de avaliação de desempenho predictivo dos algoritmos de classificação

O desempenho dos modelos de classificação foi avaliado utilizando medidas derivadas da matriz de confusão, nomeadamente acurácia, precisão, sensibilidade, especificidade e F1-score. Adicionalmente, foi utilizada a área sob a curva ROC (AUC) como medida de capacidade discriminativa do classificador. A descrição teórica dessas medidas encontra-se apresentada na Secção 2.9.2.8.

Dado que para cada par de moeda foram construídos três modelos, a fim de comparar o desempenho dos diferentes classificadores aplicados ao mesmo par de moeda, foram realizados testes estatísticos de significância a seguir descritos.

3.2.6.1 Teste de McNemar

O teste de McNemar foi aplicado para avaliar diferenças nas proporções de classificações entre dois classificadores binários avaliados sobre o mesmo conjunto de observações. Foi aplicado o teste estatístico clássico (equação 2.87).

Hipóteses:

- H_0 : não há diferença significativa nas proporções de erros discordantes entre os classificadores;
- H_1 : existe diferença significativa nas proporções de erros discordantes entre os classificadores.

Interpretação: Um valor-p inferior ao nível de significância α indica rejeição da hipótese nula, sugerindo a existência de diferença estatisticamente significativa nas proporções de erros.

3.2.6.2 Teste de DeLong

Foi aplicado o teste de DeLong para comparar as áreas sob a curva ROC (AUC) de dois classificadores avaliados sobre o mesmo conjunto de observações.

Hipóteses:

- H_0 : $AUC_1 = AUC_2$
- H_1 : $AUC_1 \neq AUC_2$

Interpretação: Um valor-p inferior ao nível de significância α indica rejeição da hipótese nula, sugerindo que existe diferença estatisticamente significativa entre as capacidades discriminativas dos classificadores.

3.2.7 Melhoramento dos modelos

Visando maximizar o desempenho predictivo e assegurar a capacidade de generalização dos modelos, realizou-se o ajuste fino dos algoritmos de machine learning. Para tal, foi aplicada a técnica de pesquisa em grelha, que consiste na avaliação sistemática de diferentes combinações de valores previamente definidos para os hiperparâmetros de cada modelo.

O processo de otimização foi conduzido em conjunto com validação cruzada do tipo KFold

adaptada para séries temporais, de forma a obter estimativas mais robustas do desempenho e reduzir o risco de sobreajuste (Packt, 2019). Para cada combinação de hiperparâmetros, o desempenho foi calculado ao longo de 10 folds temporais, sendo seleccionada a configuração que apresentou o melhor desempenho médio do AUC na validação. As combinações consideradas para cada algoritmo encontram-se na Tabela 3.4.

Tabela 3.4: Hiperparâmetros para optimização do aprendizado dos algoritmos

Algoritmo	Hiperparâmetros	EUR/USD	GBP/USD	USD/JPY	Intervalos do Grid Search	
Regressão Logística (L1)	alpha (α)	1,000	1,000	1,000	–	
	glmnet	lambda (λ)	0,01377 (0,02899)	0,00361 (0,04031)	0,00361 (0,013)	[0,00350 – 0,01400]
Árvore de Decisão	cp	0,00272 (0,00512)	0,00159 (0,01)	0,00196 (0,0314)	[0,00150 – 0,00300]	
	minsplit	38 (59)	75 (75)	96 (46)	[10 – 89]	
	rpart	maxdepth	5 (10)	11 (3)	4 (8)	[3 – 12]
	minbucket	28 (22)	17 (17)	34 (21)		[15 – 45]
	hidden	[10; 38; 54] (48)	[15; 50] (35)	[10; 49; 10] (141)		–
Redes Neurais Artificiais	algorithm	backprop	backprop	backprop		–
	threshold	21,0907 (25)	61,0907 (48)	21,0907 (39)		{21,0907, 41,0907, 61,0907}
	learningrate	0,001	0,001	0,001		{0,001; 0,01; 0,1}
	neuralnet	stepmax	10 ⁶	10 ⁶	10 ⁶	–
	rep	1	16 (1)	40 (1)		–
	err.fct	ce	ce	ce		–
	act.fct	logistic	logistic	logistic		–

Nota: Os valores a vermelho nas colunas foram utilizados para treinar os algoritmos de referência, ou seja, sem optimização. Já os outros valores apresentados nas colunas dos pares correspondem aos melhores hiperparâmetros obtidos através de pesquisa de grelha para cada par de moedas. Os outros hiperparâmetros foram mantidos para ambos casos.

Após o melhoramento dos algoritmos, procedeu-se à avaliação do desempenho predictivo com recurso à base de teste. Esta etapa tem como objectivo analisar a capacidade de generalização e robustez dos algoritmos, isto é, a sua capacidade para produzir previsões consistentes quando aplicados a dados não utilizados durante o processo de treino, tendo em conta a dependência temporal dos dados. Para tal foram aplicadas as técnicas a seguir destacadas.

3.2.7.1 Bootstrap por blocos móveis

Para estimar a variabilidade e a robustez do desempenho dos modelos, preservando a dependência temporal intrínseca aos dados, utilizou-se o método de bootstrap por blocos móveis, conforme proposto por Künsch (1989). Neste considerou-se a série temporal observada $\{X_t\}_{t=1}^T$, a qual é particionada em uma base de treinamento $\{X_t\}_{t=1}^{T_{\text{treino}}}$ e uma base de teste $\{X_t\}_{t=T_{\text{treino}}+1}^T$.

O procedimento do bootstrap por blocos móveis foi realizado exclusivamente sobre a base de treino e avaliado na base de teste, seguindo as seguintes etapas descritas a seguir:

1. **Definição do comprimento do bloco:** Escolhe-se um comprimento fixo l , correspondente ao número de observações consecutivas em cada bloco.

2. **Construção dos blocos sobrepostos:** A base de treino é decomposta em blocos sobrepostos de tamanho l , definidos explicitamente por:

$$B_1 = (X_1, \dots, X_l), \quad B_2 = (X_2, \dots, X_{l+1}), \quad \dots \quad B_{T_{\text{treino}}-l+1} = (X_{T_{\text{treino}}-l+1}, \dots, X_{T_{\text{treino}}}). \quad (3.6)$$

3. **Reamostragem dos blocos:** A partir da colecção $\{B_1, \dots, B_{T_{\text{treino}}-l+1}\}$, seleccionam-se aleatoriamente, com reposição,

$$k = \left\lceil \frac{T_{\text{treino}}}{l} \right\rceil \quad (3.7)$$

blocos.

4. **Formação da série bootstrap:** Os blocos seleccionados são concatenados na ordem sorteada, originando uma série bootstrap:

$$\{X_1^{*(b)}, X_2^{*(b)}, \dots, X_{T_{\text{treino}}}^{*(b)}\}, \quad (3.8)$$

após eventual truncamento para manter o mesmo tamanho da base de treino.

5. **Treinamento do modelo:** Em cada série bootstrap $b = 1, \dots, B$, treina-se o modelo.
6. **Avaliação do desempenho:** Para cada modelo ajustado, calcula-se a acurácia na base de teste original $(\{X_t\}_{t=T_{\text{treino}}+1}^T)$, denotada por $\hat{\theta}^{*(b)}$. A acurácia média é estimada por:

$$\hat{\theta}_{\text{média}} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}^{*(b)}. \quad (3.9)$$

O intervalo de confiança $(1 - \alpha)$ para a acurácia foi obtido por meio dos percentis da distribuição bootstrap:

$$\text{IC}_{1-\alpha} = \left[\hat{\theta}_{(\alpha/2)}^*, \hat{\theta}_{(1-\alpha/2)}^* \right]. \quad (3.10)$$

As etapas de ajuste do modelo e avaliação do desempenho são repetidas B vezes, produzindo B séries bootstrap e permitindo uma análise robusta da variabilidade e estabilidade do desempenho predictivo em presença de dependência temporal (Lahiri, 2003). Neste estudo,

o comprimento dos blocos foi fixado em $l = 30$, enquanto o número de réplicas bootstrap foi definido como $B = 500$.

3.2.7.2 Validação cruzada com janela expansiva

Para avaliar o desempenho e a robustez temporal dos modelos de previsão de forma coerente com a dependência serial das séries, utilizou-se o procedimento de validação cruzada com janela expansiva (Hyndman & Athanasopoulos 2018). A base de dados $\{X_t\}_{t=1}^T$ foi inicialmente dividida em um conjunto de treino inicial, contendo 70% das observações $\{X_t\}_{t=1}^{T_{\text{treino}}}$, e um conjunto de teste, contendo os 30% restantes $\{X_t\}_{t=T_{\text{treino}}+1}^T$, sobre o qual a validação progressiva é aplicada.

O procedimento seguiu as seguintes etapas:

1. **Inicialização da janela de treino:** Define-se a janela inicial de treino $\mathcal{T}_0$ como o conjunto de observações que compõem a base de treino original:

$$\mathcal{T}_0 = \{X_t\}_{t=1}^{T_{\text{treino}}} \quad (3.11)$$

2. **Ajuste do modelo:** O modelo é ajustado (treinado) utilizando exclusivamente as observações pertencentes à janela de treino actual $\mathcal{T}_i$.
3. **Previsão fora da amostra (*out-of-sample*):** O modelo gera previsões para as próximas k observações imediatamente subsequentes à janela de treino, definidas como a janela de validação da iteração i :

$$\mathcal{V}_i = \{X_t\}_{t=T_{\text{treino}}+(i-1)k+1}^{T_{\text{treino}}+ik} \quad (3.12)$$

4. **Expansão da janela de treino:** Na iteração $i + 1$, a janela de treino é expandida para incluir as k observações anteriormente testadas, incorporando-as ao histórico de dados conhecidos para o próximo ajuste:

$$\mathcal{T}_{i+1} = \{X_t\}_{t=1}^{T_{\text{treino}}+ik}$$

5. **Cálculo das métricas de desempenho:** Para cada iteração, calcula-se a métrica de desempenho $\hat{\theta}_i$ (a acurácia direccionada) comparando os valores previstos com os valores reais observados em $\mathcal{V}_i$. A métrica média ao longo de todas as I iterações é estimada

por:

$$\hat{\theta}_{\text{média}} = \frac{1}{I} \sum_{i=0}^{I-1} \hat{\theta}_i, \quad (3.13)$$

em que $I = (T - T_{\text{treino}})/k$ representa o número total de passos necessários para percorrer a base de teste.

6. **Intervalo de confiança:** Um intervalo de confiança $(1 - \alpha)$ para a métrica média é estimado através da distribuição empírica dos resultados obtidos em cada janela:

$$\text{IC}_{1-\alpha} = \left[\hat{\theta}_{(\alpha/2)}, \hat{\theta}_{(1-\alpha/2)} \right]. \quad (3.14)$$

3.2.8 Critérios de selecção de modelos para as estratégias de negociação

Para cada par de moedas, seleccionou-se o algoritmo de aprendizado de máquina com o desempenho mais consistente, visando sua implementação nas estratégias de negociação. A selecção dos modelos predictivos aplicados nas estratégias de negociação dos pares de moedas foi realizada com base em critérios estatísticos e financeiros. As Tabelas 3.5 e 3.6 apresentam os procedimentos adoptados para seleccionar os modelos conforme a presença ou ausência de significância estatística dos testes pareados de DeLong e McNemar.

Tabela 3.5: Critérios de selecção de modelos

Contexto	Critério	Objectivo	Interpretação
Sem significância estatística	Retorno acumulado	Avaliar desempenho financeiro	seleccionar o modelo com maior retorno acumulado
Com significância estatística	P - valor	Validar superioridade estatística	$p < 0,05$ indica que o modelo com maior acurácia é estatisticamente superior

Tabela 3.6: Procedimento de decisão para selecção de modelos

Etapa	Acção	Condição	Decisão
1	Aplicar testes estatísticos	–	Executar testes de DeLong e McNemar
2	Verificar significância	$p \leq 0,05$	seleccionar o modelo com maior acurácia ranqueada
3	Caso não haja significância	$p > 0,05$	seleccionar o modelo com maior retorno acumulado dentre os de maior acurácia

3.2.9 Estratégia de negociação

Conforme destaca Pardo (2011), um modelo de previsão, por mais acurado que seja, não constitui uma estratégia de investimento por si só, exigindo regras de execução e filtros que convertam sinais em operações de mercado. Sob essa óptica, a estratégia de negociação implementada transforma as previsões probabilísticas dos modelos em decisões operacionais, integrando filtros técnicos e um mecanismo de controle de risco baseado no valor em risco (VaR). A estratégia opera de forma condicional, os sinais de execução dependem tanto dos limites críticos de probabilidade (U_Prob e L_Prob) quanto de filtros de exaustão e persistência, conforme detalhado no Algoritmo 2.

A estratégia utiliza o índice de força relativa (IFR) e o expoente de Hurst (H) como critérios de validação. O IFR actua na identificação de zonas de sobrevenida para posições compradas, enquanto o expoente de Hurst filtra séries com características de reversão à média.

Para a gestão de risco, utilizou-se um *stop loss* baseado no VaR, além de um critério de encerramento temporal ao fim de cada sessão diária (Secção_Fechada). Aplicou-se também custos operacionais. A tabela 4.16 apresenta esses parâmetros operacionais da estratégia para cada par de moeda.

Algorithm 2: Regra de Negociação Baseada nas Probabilidades dos Modelos e Filtros**Entrada:**

- Indicadores técnicos e seus periodos: $IFR5_i$, $Hurst15_i$
- Cutoff: $U.Prob$, $L.Prob$, I , $H(u)$, $H(d)$
- Stop-loss: VaR
- Preços: P_{open} (preço de abertura), $P_{current}$ (preço actual)
- Controlo da sessão: Sessão Fechada (booleano)

Saída:

- Retorno acumulado

```

for Cada incremento de tempo  $i$  nos dados do mercado do
  // Calcular a probabilidade predicta pelo modelo
   $prob \leftarrow P(Y = 1 | X_i)$ 
  if  $prob > U.Prob$  and  $((IFR15_i < I) \text{ or } (Hurst15_i < H(u)))$  then
    ABRIR posição de COMPRA no preço  $P_{open}$ 
    // Monitorar posição até stop-loss ou fechamento da sessão
    repeat
       $R_{buy} \leftarrow (P_{current}/P_{open}) - 1$ 
    until  $R_{buy} \leq -VaR$  or (Secção Fechada = TRUE)
    FECHAR posição de COMPRA
  end
  else if  $prob < L.Prob$  and  $(Hurst15_i < H(d))$  then
    ABRIR posição de VENDA no preço  $P_{open}$ 
    // Monitorar posição até stop-loss ou fechamento da sessão
    repeat
       $R_{sell} \leftarrow (P_{current}/P_{open}) - 1$ 
    until  $R_{sell} \leq -VaR$  or (Secção Fechada = TRUE)
    FECHAR posição de VENDA
  end
  else
    SEM posição
  end
end

```

3.2.10 Análise de risco e desempenho financeiro

A análise de desempenho e de risco financeiro dos algoritmos e da estratégia de negociação foi conduzida por meio de um conjunto de medidas de risco e desempenho. As medidas consideradas incluem curtose, assimetria, desvio padrão negativo, volatilidade (desvio padrão), volatilidade anualizada, valor em risco (VaR), valor em risco condicional (CVaR), perda máxima acumulada, índice de Sharpe e índice de Sortino.

Essas medidas foram calculadas com base nos resultados observados na base fora da amostra, permitindo uma comparação directa entre o desempenho do modelo de previsão isolado e o desempenho da estratégia de negociação implementada. A análise conjunta dessas medidas possibilitou avaliar não apenas a capacidade de geração de retorno, mas também o risco assumido e a eficiência risco-retorno das abordagens analisadas.

Capítulo 4

RESULTADOS E DISCUSSÃO

4.1 Análise exploratória

A Figura 4.1 apresenta os gráficos dos preços de fechamento e o comportamento dos retornos para os pares de moedas seleccionados (EUR/USD, USD/JPY e GBP/USD). Subsequentemente, apresenta-se a análise descritiva da estrutura temporal dos preços e da distribuição dos retornos diários (Tabela 4.1 e Figura 4.2). Esta análise compreende o intervalo de 4 de fevereiro de 2013 a 31 de janeiro de 2025.

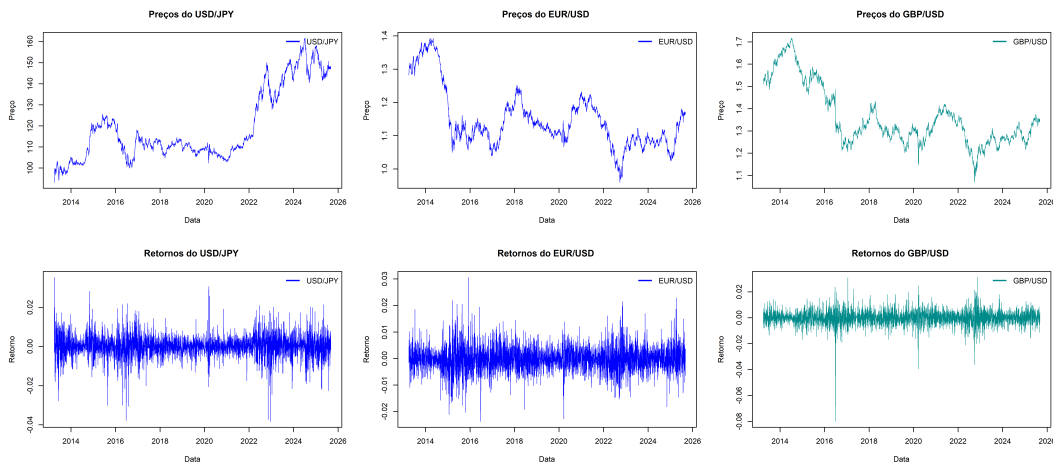


Figura 4.1: Preços de fechamento e retornos aritméticos para os pares.

Tabela 4.1: Estatísticas Descritivas

Par	Mínima	Q1	Mediana	Q3	Máxima	Média	CV
USD/JPY	75,674	104,858	111,237	120,459	161,665	112,816	0,15
EUR/USD	0,8267	1,0869	1,1667	1,2934	1,5997	1,1819	0,13
GBP/USD	1,0683	1,3425	1,5512	1,65	2,1065	1,5366	0,14

Na Figura 4.1, observa-se que o par USD/JPY apresenta um regime de alta acentuado após 2020. Em contraste, os pares EUR/USD e GBP/USD evidenciam trajetórias de baixa, intercaladas por recuperações pontuais. Nos gráficos de retornos aritméticos, nota-se que os valores oscilam em torno de zero, com picos associados a eventos extremos em 2016, 2020 e 2022.

A Tabela 4.1 apresenta as estatísticas descritivas dos preços de fechamento dos pares. O coeficiente de variação (CV) evidencia a dispersão relativa dos dados em relação à média de cada série. Observa-se que o USD/JPY apresenta a maior variabilidade relativa entre os pares analisados (0,15), seguido pelo GBP/USD e pelo EUR/USD. Indicando um comportamento de variação média ao longo do período.

Na Figura 4.2, observa-se a distribuição dos retornos diários dos pares USD/JPY, EUR/USD e GBP/USD. Em todos os casos, os histogramas evidenciam um formato leptocúrtico, isto é, mais afilado e com caudas mais pesadas em comparação à distribuição normal.

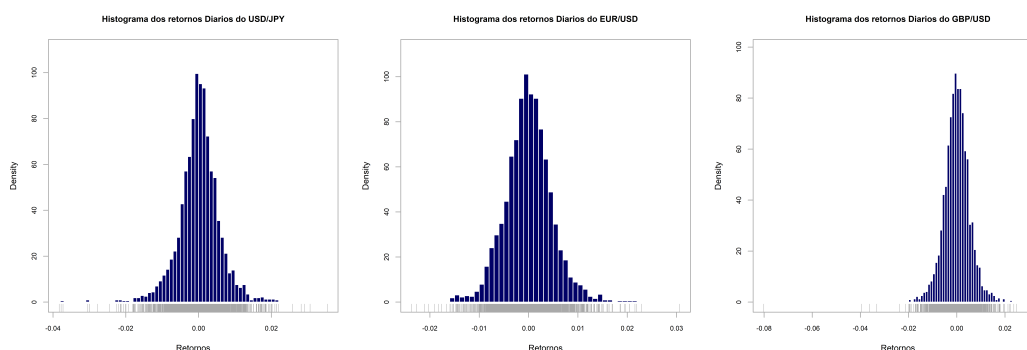


Figura 4.2: Histograma dos retornos aritméticos dos pares de moeda.

Mapa de calor de correlação entre os activos

O mapa de calor de correlação dos activos evidencia relações relevantes para compreender a estrutura intermercados entre os retornos no mercado cambial e outros mercados. Observa-se que o índice do dólar norte-americano (DXY) apresenta forte correlação negativa com EUR/USD (-0,96) e GBP/USD (-0,72), enquanto a correlação com USD/JPY é positiva moderada (0,57), reflectindo a influência central do dólar sobre os principais pares. Entre as moedas, a correlação positiva moderada entre EUR/USD e GBP/USD (0,62) indica que ambas tendem a reagir de forma análoga ao dólar, ao passo que o USD/JPY mantém correlações negativas com os pares europeus.

As correlações entre pares de moedas e índices de acções (Euro Stoxx 50, Nikkei 225, S&P500) situam-se em níveis baixos a moderados, destacando-se a relação positiva entre USD/JPY

e Nikkei 225 (0,53). Entre os índices, observam-se correlações fortes, nomeadamente Euro Stoxx 50 com S&P500 (0,80) e com Nikkei 225 (0,73). O petróleo (WTI - West Texas Intermediate) apresenta correlações fracas com moedas e os índices (Figura: 4.3).

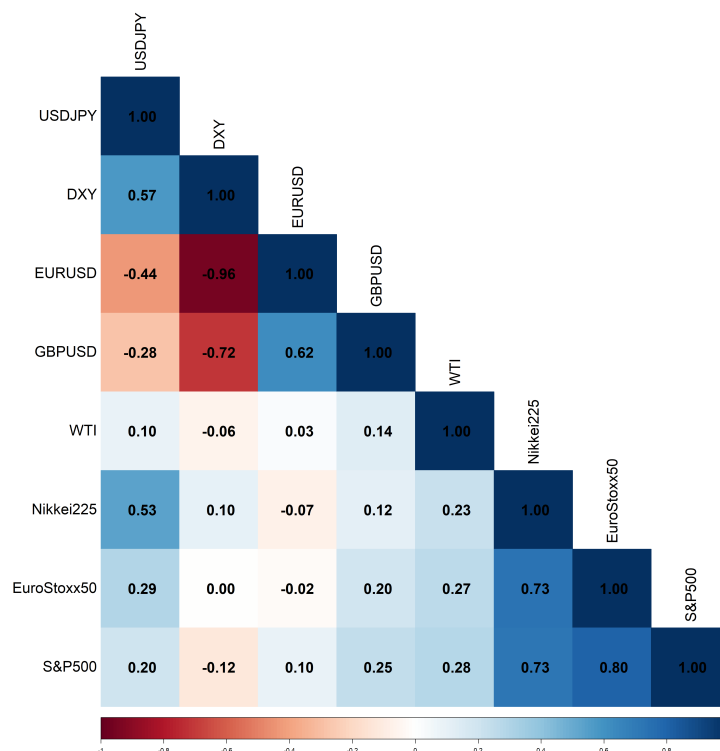


Figura 4.3: Heat Map da correlação entre os retornos diários dos activos.

Análise de dispersão

A Figura 4.4 apresenta os gráficos de dispersão do USD/JPY, visando avaliar a relação entre os retornos actuais e os do dia subsequente sob diferentes condições dos indicadores IFR e Oscilador Estocástico. No IFR de 15 períodos, observa-se que, em níveis de sobrecompra (≥ 70), os valores concentram-se em torno de zero com inclinação negativa da recta, sugerindo um esgotamento do potencial de ganhos. Acima de 80, embora a concentração seja menor, a inclinação negativa persiste. Em níveis de sobrevenda (abaixo de 30 e 20), a dispersão acentua-se em valores negativos dos retornos, mas a inclinação positiva da regressão indica uma possível recuperação dos preços. Já o oscilador estocástico, para a mesma periodicidade, revela relações marginais com rectas praticamente nulas, evidenciando sua reduzida relação com o retorno do alvo.

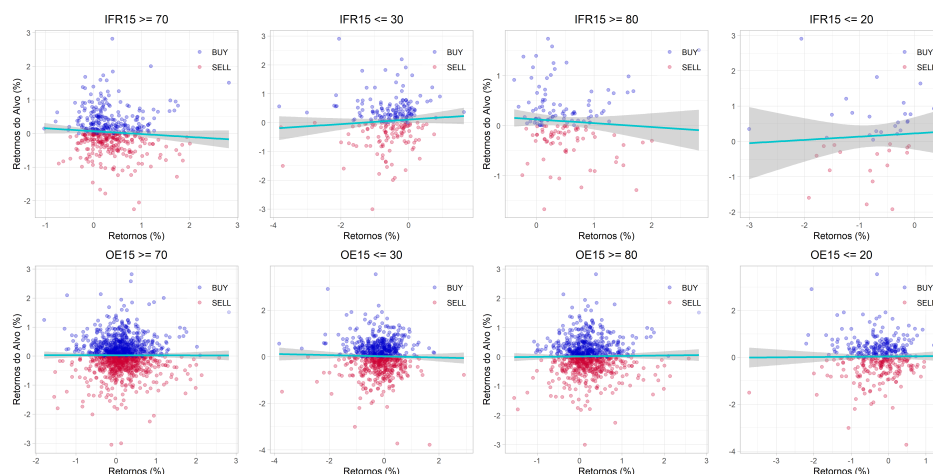


Figura 4.4: Relação entre retornos e retornos do alvo sob diferentes condições do índice de força relativa e oscilador estocástico do USD/JPY.

No gráfico de dispersão do EUR/USD (Figura 4.5), observa-se que, em sobrecompra ($IFR \geq 70$), os retornos mantêm-se concentrados em torno de zero, com leve inclinação positiva da recta. Quando o IFR15 atinge níveis extremos (acima de 80), a inclinação torna-se negativa, sugerindo que movimentos positivos tendem a perder força e sinalizando possível reversão no dia subsequente. Já em contextos de sobrevenda ($IFR15 < 30$ ou 20), verifica-se inclinação positiva das rectas, efeito mais evidente abaixo de 20. Esse padrão indica potencial de recuperação.

Nos resultados do oscilador estocástico (OE15), as rectas apresentam relação linear praticamente nula entre retornos correntes e futuros. Verifica-se uma fraca inclinação positiva em níveis de sobrecompra extremos ($OE15 \geq 80$) (Figura 4.5).

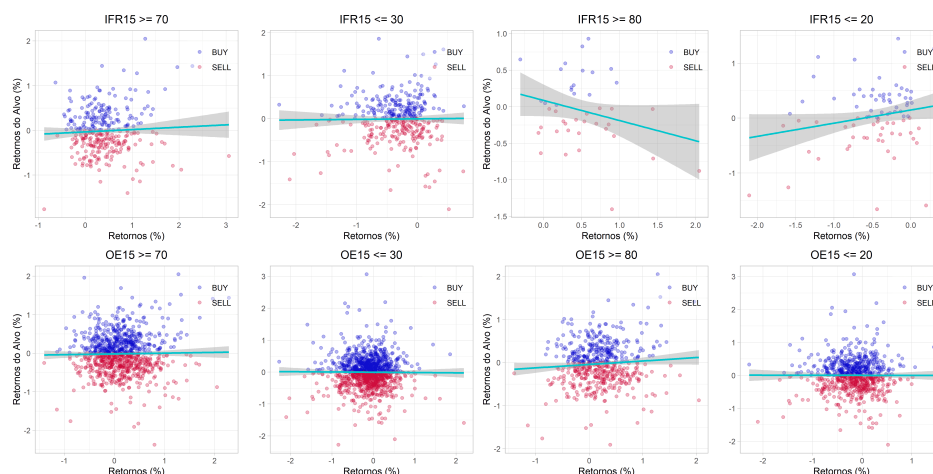


Figura 4.5: Relação entre retornos e retornos do alvo sob diferentes condições do índice de força relativa e oscilador estocástico do EUR/USD.

A Figura 4.6 apresenta os gráficos de dispersão dos retornos do GBP/USD. Observa-se que, em sobrecompra ($IFR \geq 70$ e 80), as rectas assumem inclinação negativa, o que sugere uma possível exaustão da tendência de alta e sinaliza potencial reversão nos retornos futuros. Em contrapartida, a sobrevenda ($IFR \leq 30$ e 20), destaca-se para retornos negativos no dia corrente acompanhados de inclinação positiva, apontando para uma possível recuperação, isto é, uma alta no dia subsequente. Quanto ao oscilador estocástico, os resultados mostram-se ligeiramente mais expressivos que nos pares USD/JPY e EUR/USD. Na condição de sobrevenda (≤ 30 e 20), as rectas apresentam inclinação positiva leve, indicando relação linear entre retornos correntes e futuros.

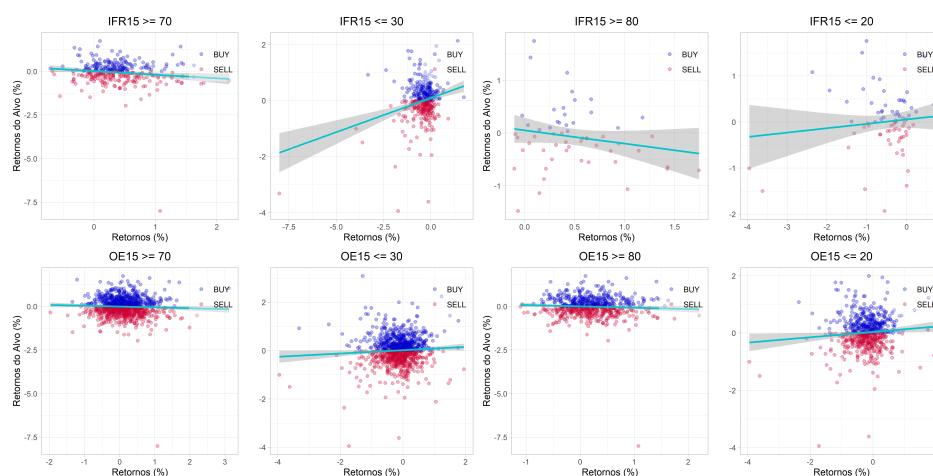


Figura 4.6: Relação entre retornos e retornos do alvo sob diferentes condições do índice de força relativa e oscilador estocástico do GBP/USD.

Análise gráfica dos retornos médios por activo e horizonte temporal

A Figura 4.7 apresenta os retornos médios por activo e período. Observa-se que, à excepção do USD/JPY, as taxas de variação dos pares cambiais estudados mantêm-se próximas de zero e negativas ao longo do tempo, indicando um possível comportamento de reversão à média. No par USD/JPY, o aumento do número de períodos na taxa de variação revela retornos médios positivos, indicando um comportamento de tendência. Em contraste, os índices e o WTI apresentam médias positivas superiores com o aumento do horizonte temporal. Destaca-se, mais uma vez, a relação inversa do GBP/USD e do EUR/USD com o índice DXY.

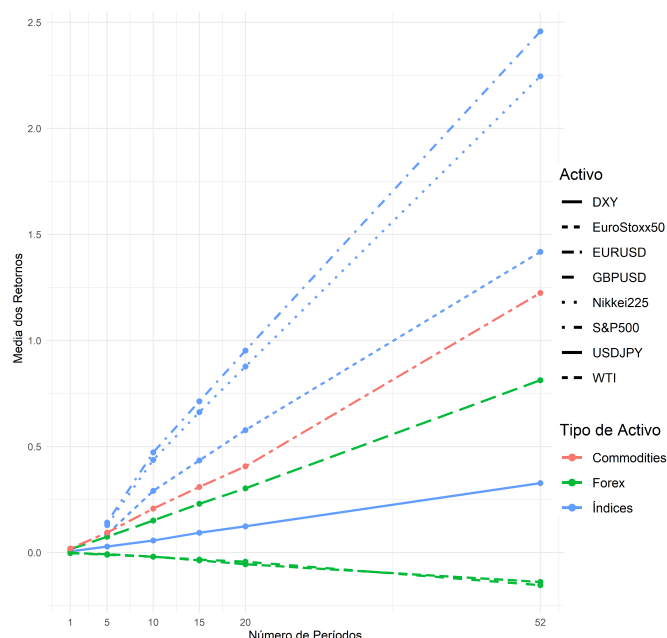


Figura 4.7: Retornos médios por activo e periodo.

Evolução temporal da estacionaridade e do coeficiente de Hurst

As Figuras 4.37, 4.38 e 4.39 em anexo, apresentam a evolução temporal dos p-valores do teste ADF em cada par, revelando que os retornos não são constantemente estacionários. Em horizontes de 10 e 15 períodos, verificam-se oscilações acentuadas em valores acima de 0,05, o que reflecte um comportamento não estacionário predominante. Já em janelas mais longas, de forma mais acentuada no *PER52*, observa-se que intervalos prolongados de comportamento estacionário são interrompidos por picos significativos nos p-valores, o que sugere possíveis mudanças de regime de mercado. Esta evidência aponta que os mercados tendem a alternar entre fases de relativa previsibilidade e de elevada aleatoriedade nesse periodo.

No que se refere à evolução temporal do coeficiente de Hurst nos pares de moedas, observa-se (Figuras 4.34, 4.35 e 4.36 em anexo) que, para horizontes de curto prazo, como 10 e 15 períodos, os valores de H situam-se maioritariamente abaixo de 0,5. Tal comportamento indica a predominância de um regime de antipersistência, no qual as séries apresentam reversão à média, movimentos de alta tendem a ser sucedidos por movimentos de baixa, e vice-versa. Em contrapartida, ao analisar janelas mais prolongados, como 20 e 52 períodos, verifica-se maior oscilação em torno do valor central (0,5). Essa alternância sugere que o mercado intercala períodos de persistência ($H > 0,5$), caracterizados por movimentos tendenciais, com intervalos de antipersistência ($H < 0,5$), evidenciando uma dinâmica de mercado mais heterogênea em escalas temporais maiores.

4.2 Selecção de variáveis

Nesta secção, apresentam-se as variáveis seleccionadas, baseado no critério de selecção apresentado no algoritmo 1, dado que verificou-se uma forte correlação entre várias das 209 variáveis independentes (Figuras 4.52, 4.53 e 4.54 em anexo).

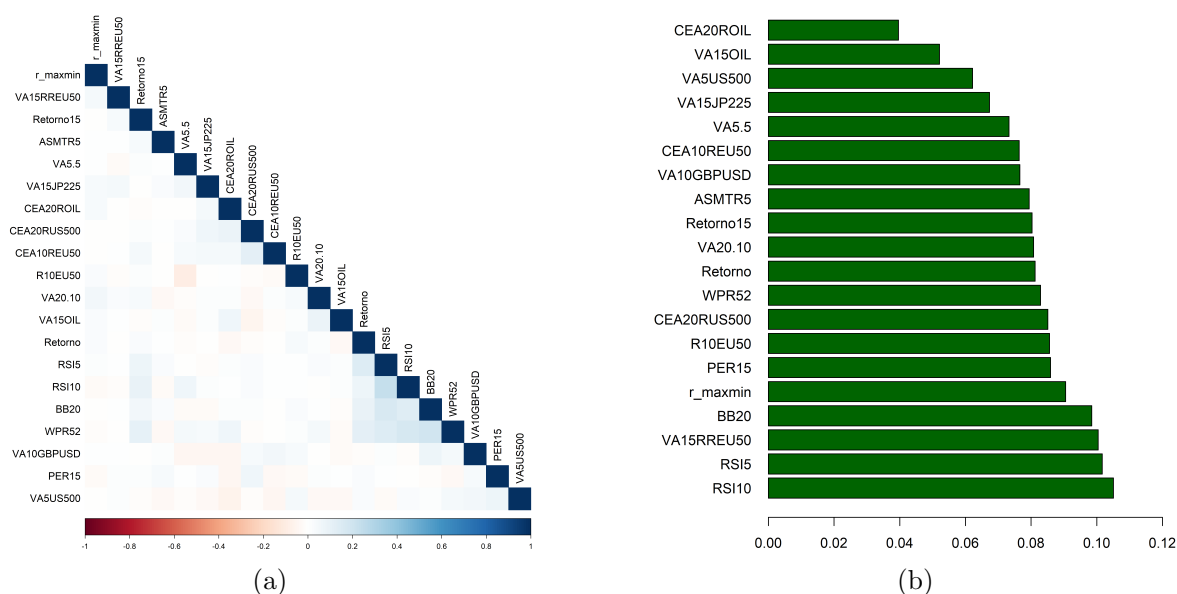


Figura 4.8: EUR/USD: (a) Correlação entre as variáveis independentes seleccionadas; (b) Correlação entre as variáveis independentes seleccionadas e o alvo

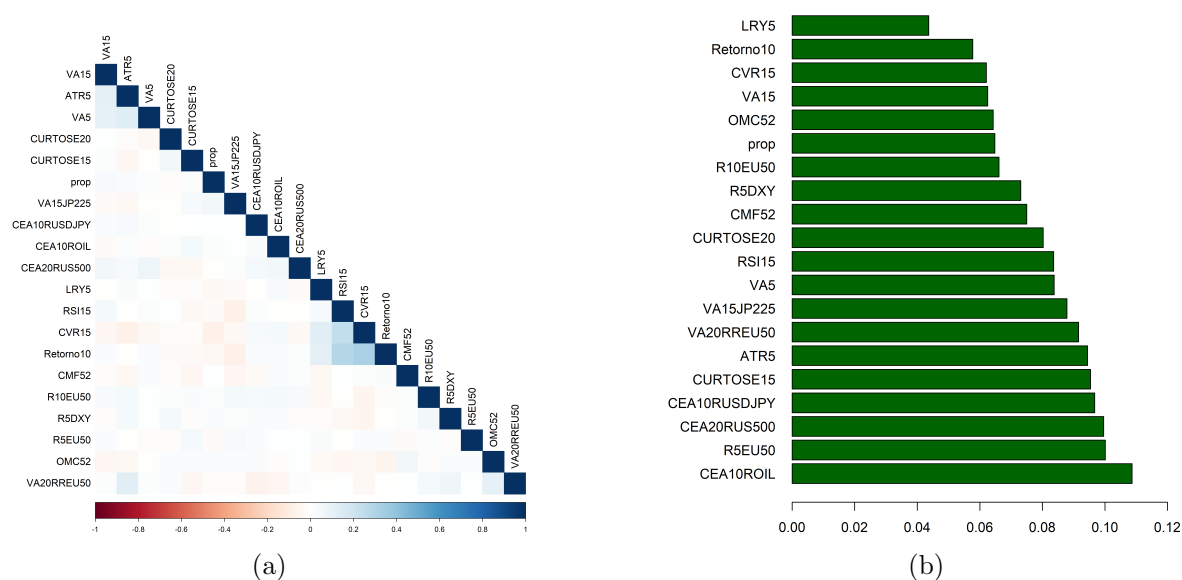


Figura 4.9: GBP/USD: (a) Correlação entre as variáveis independentes seleccionadas; (b) Correlação entre as variáveis independentes e o alvo

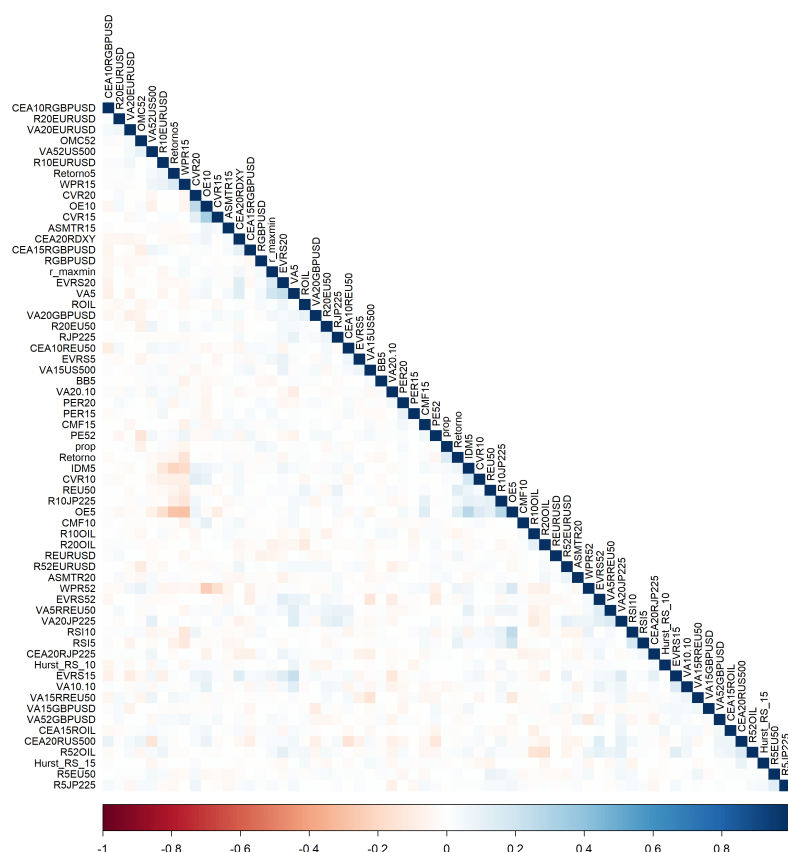


Figura 4.10: Correlação entre as variáveis independentes seleccionadas para os modelos do USD/JPY

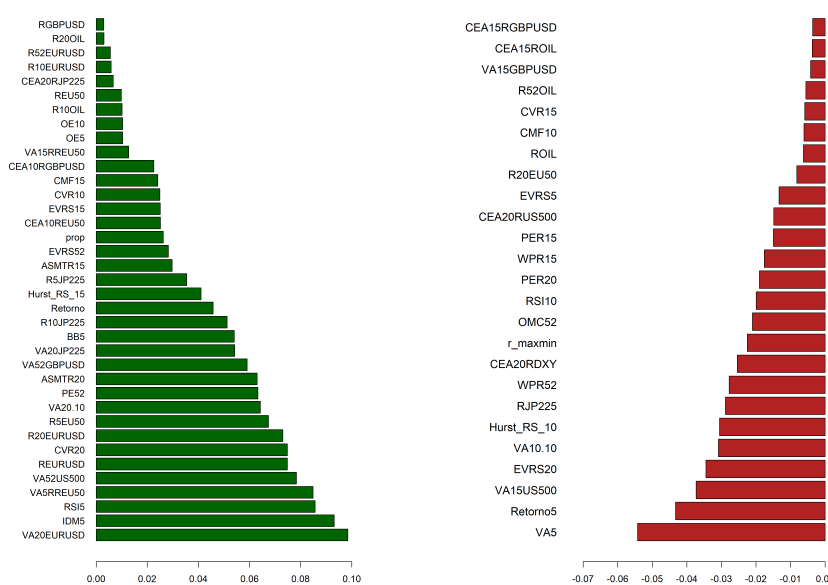


Figura 4.11: Correlações positivas e negativas entre as variáveis seleccionadas para os modelos do USD/JPY x Alvo

4.3 Resultados de desempenho geral dos algoritmos

Esta secção apresenta os resultados das previsões dos algoritmos nos dados fora da amostra. A análise do desempenho dos modelos é feita através de métodos estatísticos e por tipo de algoritmo, os quais foram discutidos nas secções anteriores (Secção 3.2).

4.3.1 Análise do desempenho geral dos algoritmos

Tabela 4.2: Desempenho geral dos algoritmos na previsão da direcção dos retornos diários.

Par	Algoritmo	Acurácia	Precisão	Recall	Especificidade	F1-Score	AUC
USD/JPY	RL - L	66,75 (63,09)	70,62 (68,00)	71,26 (66,44)	60,79 (58,66)	70,94 (67,21)	70,00 (66,38)
	AD - C	57,33 (51,31)	61,52 (57,08)	66,90 (58,39)	44,68 (41,95)	64,10 (57,73)	54,00 (50,17)
	RNA - BP	61,78 (58,64)	58,16 (56,09)	69,70 (66,12)	54,61 (51,65)	63,41 (60,70)	62,00 (60,90)
GBP/USD	RL - L	64,00 (57,59)	60,72 (57,07)	62,01 (58,58)	60,52 (56,62)	61,36 (57,81)	62,23 (59,20)
	AD - C	54,71 (52,62)	55,15 (53,70)	54,25 (43,38)	55,15 (62,01)	54,71 (47,99)	55,00 (55,22)
	RNA - BP	60,08 (51,18)	56,62 (54,03)	61,24 (51,49)	59,07 (50,83)	58,84 (52,72)	60,00 (50,06)
EUR/USD	RL - L	59,95 (59,16)	64,33 (62,65)	51,01 (52,53)	69,57 (66,30)	56,90 (57,14)	66,00 (63,70)
	AD - C	57,20 (54,32)	58,04 (52,36)	40,22 (57,34)	72,98 (51,52)	47,51 (54,73)	58,00 (55,64)
	RNA - BP	61,39 (53,01)	66,03 (57,07)	58,84 (51,09)	64,39 (55,24)	62,23 (53,92)	65,00 (55,90)

A Tabela 4.2 apresenta o desempenho dos algoritmos fora da amostra, com e sem a optimização. Os valores principais apresentados na tabela correspondem aos modelos optimizados por pesquisa de grelha, enquanto os valores apresentados entre parênteses referem-se aos resultados obtidos sem optimização de hiperparâmetros. De forma geral, observa-se que os algoritmos com optimização apresentam capacidade predictiva superior.

No que se refere aos modelos optimizados, a regressão logística penalizada via Lasso (RL-L) destacou-se de forma consistente, apresentando valores equilibrados de acurácia, precisão, recall, especificidade e F1-score nos três pares de moedas. Para o par USD/JPY, obteve-se acurácia de 66,75%, precisão de 70,62%, recall de 70,62%, especificidade de 60,79% e F1-score de 70,94%. Resultados semelhantes foram observados para os pares GBP/USD (61,26% de acurácia, 60,72% de precisão, 62,01% de recall, 69,52% de especificidade e 61,36% de F1-score) e EUR/USD (59,95% de acurácia, 64,33% de precisão, 51,01% de recall, 69,57% de especificidade e 56,90% de F1-score). Esses resultados demonstram que o modelo conseguiu não apenas classificar correctamente uma proporção significativa de observações, mas também manter uma relativa boa taxa de acertos tanto nos eventos positivos (Compra) quanto nos negativos (Venda).

Seguindo a análise anterior, observa-se que, para o par EUR/USD, o recall (sensibilidade) da RL-L foi de 51,01%, indicando que o modelo identificou correctamente aproximadamente metade das oportunidades reais de compra. Esse desempenho é considerado reduzido, uma vez que implica uma taxa elevada de falsos negativos (oportunidades de compra não aproveitadas). No entanto, tal limitação é parcialmente mitigada pela precisão de 64,33%, a qual

demonstra que, quando o modelo emite um sinal de compra, apresenta uma assertividade satisfatória.

Por outro lado, o algoritmo AD-C (árvore de decisão para classificação) apresentou desempenho inferior em todos os pares de moedas, com medidas mais baixas e menos consistentes. No par GBP/USD, por exemplo, a acurácia foi de 54,71%, acompanhada de precisão de 55,15%, recall de 54,25%, especificidade de 55,15% e F1-score de 54,71. Tais resultados evidenciam a dificuldade do modelo em discriminar as classes de forma eficaz. Os pares EUR/USD e USD/JPY seguiram essa tendência, com acurácias de 57,20% e 57,33%, respectivamente.

As redes neurais artificiais via retropropagação (RNA-BP) demonstraram competitividade em relação à regressão logística penalizada, com resultados relativamente próximos. O algoritmo atingiu acurácias de 61,39% no EUR/USD, 60,08% no GBP/USD e 61,78% no USD/JPY.

Observa-se que, para os três pares de moedas analisados, os modelos otimizados tendem a apresentar um AUC superior quando comparados com as versões sem otimização. Os valores de AUC obtidos situam-se majoritariamente acima ou próximos de 0.60, indicando que os modelos apresentam capacidade discriminatória superior ao nível aleatório.

Assim, considerando o desempenho superior observado nos dados fora da amostra, as análises subsequentes deste estudo concentram-se apenas nos modelos otimizados, por estes apresentarem resultados globais mais favoráveis na previsão da direção dos retornos diários dos pares cambiais analisados.

4.3.2 Matriz de Confusão

As matrizes de confusão (Figuras 4.12, 4.13 e 4.14) ilustram o desempenho dos modelos na classificação das observações. Observa-se que, nos pares EUR/USD e USD/JPY, os algoritmos baseados em árvore de decisão tendem a classificar com maior assertividade apenas uma das categorias, onde o EUR/USD apresenta melhor desempenho na identificação de vendas, enquanto o USD/JPY classifica melhor as compras. Em contrapartida, as matrizes da regressão logística e das redes neurais artificiais exibem maior equilíbrio entre as classes. Quanto ao par GBP/USD, as matrizes demonstram relativa estabilidade na distribuição das classificações tanto para posições de venda quanto de compra.

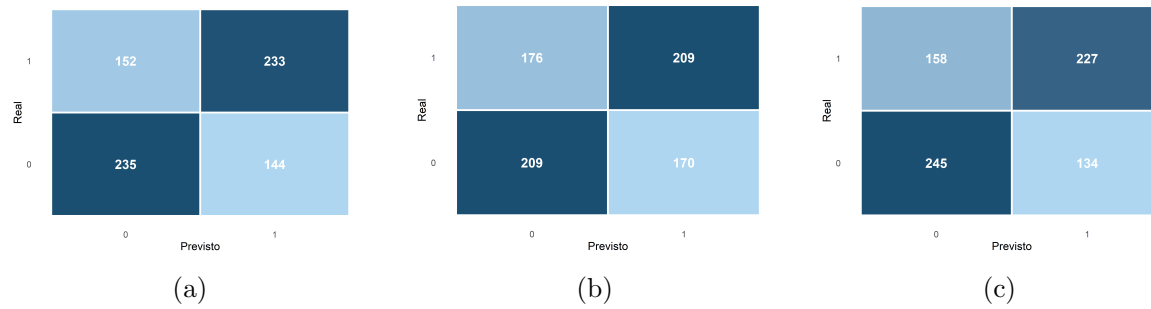


Figura 4.12: Matrizes de confusão dos modelos do GBP/USD: (a) Regressão logística penalizada via L1; (b) Árvore de decisão; (c) Rede neural artificial.

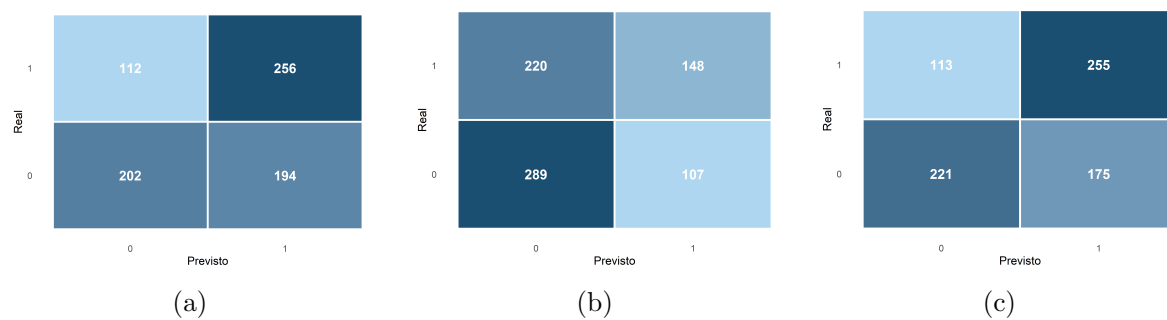


Figura 4.13: Matrizes de confusão dos modelos do EUR/USD: (a) Regressão logística penalizada via L1; (b) Árvore de decisão; (c) Rede neural artificial.

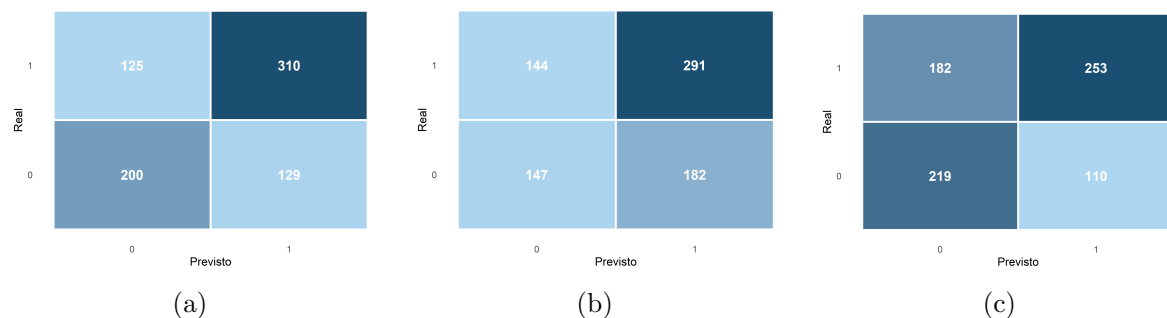


Figura 4.14: Matrizes de confusão dos modelos do USD/JPY: (a) Regressão logística penalizada via L1; (b) Árvore de decisão; (c) Rede neural artificial.

4.3.3 Curvas ROC

Dando continuidade as análises, as curvas ROC (Figuras 4.16, 4.15 e 4.17) ilustram visualmente a capacidade discriminatória dos modelos otimizados ao longo de diferentes limiares de classificação.

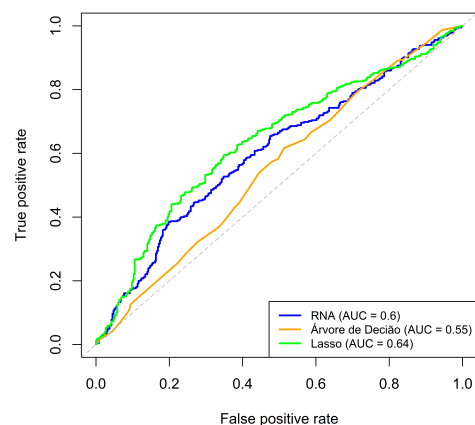


Figura 4.15: Curvas ROC dos modelos do GBP/USD.

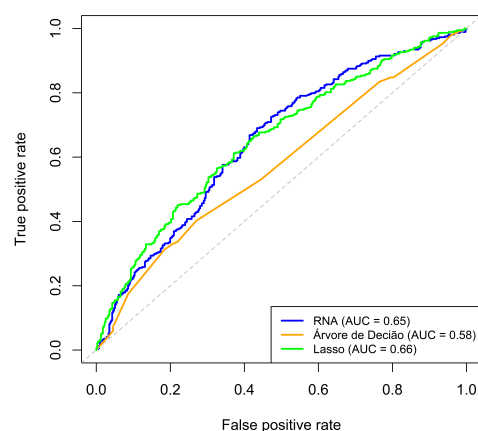


Figura 4.16: Curvas ROC dos modelos do EUR/USD.

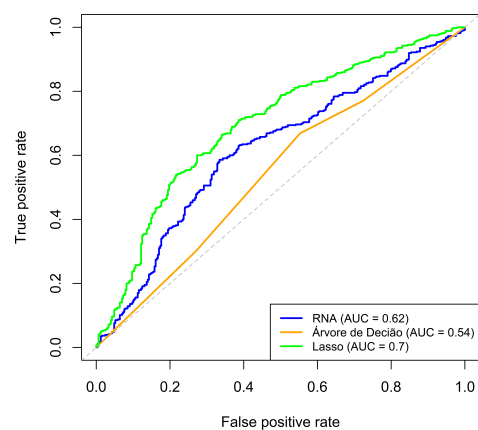


Figura 4.17: Curvas ROC dos modelos do USD/JPY.

No par EUR/USD, a regressão logística penalizada via lasso apresentou uma AUC de aproximadamente 0,66, ligeiramente superior à da rede neural artificial (0,65). Observa-se, contudo, que quando a Taxa de Falsos Positivos excede 0,4, a Taxa de Verdadeiros Positivos

da RNA torna-se ligeiramente superior à do lasso; inversamente, abaixo desse patamar, o lasso apresenta melhor desempenho. Ambos os modelos situam-se consideravelmente acima da árvore de decisão (0,58). No par GBP/USD, o lasso e a RNA obtiveram valores de 0,64 e 0,60, respectivamente, enquanto a árvore de decisão manteve um desempenho inferior (0,55). Para o par USD/JPY, constatou-se um desempenho de 0,70 na regressão logística, seguido pela RNA com 0,62 e pela árvore de decisão com 0,57.

4.3.4 Comparação entre os algoritmos

Observou-se que as medidas de desempenho dos modelos finais (Tabela 4.2) apresentam valores semelhantes. Torna-se, portanto, pertinente avaliar se tais semelhanças possuem significância estatística.

Par EUR/USD

Tabela 4.3: Testes de McNemar pareado entre os algoritmos.

Algoritmo	RL - Lasso	AD - C	RNA - BP
RL - L	-	$X^2=4,83$ (p=0,03)	$X^2=2,73$ (p=0,1)
AD - C		-	$X^2=1,39$ (p=0,24)
RNA - BP			-

Na comparação pareada dos classificadores (Tabela 4.3), observa-se, por meio do valor-p, que o algoritmo RL-L difere significativamente da árvore de decisão (0,03), mas não apresenta diferença estatística em relação à rede neural (0,10). Por outro lado, a comparação entre a árvore de decisão e a rede neural não revelou diferença significativa (0,24), sugerindo que, embora a árvore apresente medidas inferiores, seu desempenho não é estatisticamente distinto do observado para a rede.

Tabela 4.4: Testes de Delong pareado entre os algoritmos.

Algoritmo	RL - Lasso	AD - C	RNA - BP
RL - L	-	$Z=3,4$ (p=0,000)	$Z=-0,37$ (p=0,71)
AD - C		-	$Z=-3,79$ (p=0,00)
RNA - BP			-

O teste de DeLong (Tabela 4.4) confirma que a curva ROC do algoritmo RL-L e RNA - BP apresentam capacidade discriminativa significativamente diferente ($p < 0,05$) ao da árvore de decisão. Em contrapartida, a diferença entre o lasso e a rede neural artificial não é estatisticamente significativa (0,71).

Em síntese, o algoritmo RNA-BP apresentou ligeira vantagem na acurácia (61,39%) em relação ao lasso (59,95%), porém, tal superioridade não é estatisticamente significativa.

Par GBP/USD

Na comparação pareada do par GBP/USD (Tabela 4.5), verifica-se, pelos valores-p, que o algoritmo RL-L difere significativamente do algoritmo AD-C (0,00), mas não apresenta diferença estatística em relação à RNA-BP (0,80). Além disso, a AD-C apresenta desempenho estatisticamente distinto da RNA-BP (0,00).

Tabela 4.5: Testes de McNemar pareado para os algoritmos.

Algoritmo	RL - Lasso	AD - C	RNA - BP
RL - Lasso	-	$X^2=9,30$ (p=0,00)	$X^2=0,06$ (p=0,80)
AD - C		-	$X^2=8,11$ (p=0,00)
RNA - BP			-

De forma coerente, o valor-p do teste de DeLong (Tabela 4.6) confirma o maior poder discriminativo da regressão logística penalizada em relação à árvore de decisão (0,00), não havendo diferença significativa em comparação à rede neural (0,97). A rede neural, por sua vez, é significativamente distinta da árvore de decisão (0,00). Assim, os algoritmos RL-L e RNA-BP consolidam-se como algoritmos com desempenho similar para o par GBP/USD.

Tabela 4.6: Testes de Delong pareado para os algoritmos.

Algoritmo	RL - Lasso	AD - C	RNA - BP
RL - Lasso	-	$Z=3,50$ (p=0,00)	$Z=-0,04$ (p=0,97)
AD - C		-	$Z=-3,57$ (p=0,00)
RNA - BP			-

Relacionando as acurácias (Tabela 4.2) e as tabelas 4.6 e 4.5, o algoritmo RL-L manteve a maior acurácia (61,26%), se comparado a rede neural (60,08%), porém, confirmou-se que essa superioridade é não significativa, pelo que os testes pareados indicam que o desempenho não é estatisticamente diferente, demonstrando competitividade entre ambos classificadores.

Par JPY/USD

Na análise pareada realizada por meio do teste de McNemar para o par USD/JPY (Tabela 4.7), observa-se, pelos valores-p, que o algoritmo RL-L não difere estatisticamente da AD-C (0,07), mas apresenta diferença significativa em relação à RNA (0,02). Adicionalmente, a rede neural artificial é estatisticamente distinta da árvore de decisão (0,00).

Tabela 4.7: Testes de McNemar pareado para os algoritmos.

Algoritmo	RL - Lasso	AD - C	RNA - BP
RL - Lasso	-	$X^2=3,32$ (p=0,07)	$X^2=5,47$ (p=0,02)
AD - C		-	$X^2=15,95$ (p=0,00)
RNA - BP			-

Os valores-p do teste de DeLong para o par USD/JPY (Tabela 4.8) indicam que todos os algoritmos apresentam capacidades discriminativas significativamente distintas entre si ($p < 0,05$). Esse resultado confirma que as diferenças de desempenho observadas na Tabela 4.2 são estatisticamente relevantes, sugerindo que os modelos capturam padrões distintos nos dados. Nesse cenário, a regressão logística (Tabela 4.2) destacou-se como o modelo de melhor desempenho, apresentando a maior acurácia entre os algoritmos comparados.

Tabela 4.8: Testes de Delong pareado para os algoritmos.

Algoritmo	RL - Lasso	AD - C	RNA - BP
RL - Lasso	-	Z=2,68 (p=0,01)	Z=-3,58 (p=0,00)
AD - C		-	Z=-5,72 (p=0,00)
RNA - BP			-

4.4 Avaliação do desempenho dos algoritmos utilizando bootstrap por blocos moveis

A Tabela 4.9 apresenta os resultados nos dados fora da amostra obtidos por meio do bootstrap móvel por blocos para o par EUR/USD. Observa-se que a regressão logística penalizada via lasso (RL-L) alcança a maior acurácia média (59,58%), seguido das redes neurais (56,54%), e da árvore de decisão (54,01%). Observa-se que as acurácias dos algoritmos estáticos, 59,95%, 57,20% e 61,39% (Tabela 4.2), situam-se dentro dos respectivos intervalos de confiança (95%) do bootstrap. A Figura 4.49 ilustra a distribuição das acurácias dos respectivos algoritmos, por meio de boxplots.

Tabela 4.9: Avaliação do desempenho dos modelos do EURUSD utilizando bootstrap por blocos moveis.

Algoritmo	Acurácia Média	IC (2,5%)	IC (97,5%)
RL - L	59,58%	57,19%	61,91
AD - C	54,01%	50,39%	57,32
RNA - BP	56,54%	55,39%	61,67

Para o par GBP/USD, conforme indicado na Tabela 4.10, a regressão logística penalizada via lasso mantém o melhor desempenho predictivo, com acurácia média de 60,77%, seguido das redes neurais que apresentam acurácia de 57,21. A árvore de decisão apresenta acurácia média de 52,93%, ao longo das reamostragens. Observa-se que as acurácias dos algoritmos estáticos, 61,26%, 54,71% e 60,08% (Tabela 4.2), situam-se dentro dos respectivos intervalos de confiança (95%) do bootstrap. Os boxplots apresentados na Figura 4.50 mostram a distribuição das acurácias das reamostragens para cada respectivo algoritmo no par GBP/USD.

Tabela 4.10: Avaliação do desempenho dos modelos do GBPUSD utilizando bootstrap por blocos moveis (%).

Algoritmo	Acurácia Média	IC (2,5%)	IC (97,5%)
RL - L	60,77%	59,03%	62,43
AD - C	52,93%	49,53%	56,15
RNA - BP	57,21%	51,83%	60,67

Para o par USD/JPY, os resultados do bootstrap móvel por blocos (Tabela 4.11) indicam que o algoritmo RL-L atinge a maior acurácia média (63,71%), seguido das redes neurais (58,77%), e da árvore de decisão (52,14%). Observa-se as acurácias dos modelos estáticos, 66,75%, 57,33% e 61,78% (Tabela 4.2), situam-se ligeiramente acima dos limites superiores dos respectivos intervalos de confiança de 95%. Os boxplots da Figura 4.51 apresenta o comportamento das distribuições de acurácia entre os algoritmos analisados.

Tabela 4.11: Avaliação do desempenho dos modelos do USDJPY utilizando bootstrap por blocos moveis (%)

Algoritmo	Acurácia Média	IC (2,5%)	IC (97,5%)
RL - L	63,71%	60,86%	66,09
AD - C	52,14%	47,25%	56,35
RNA - BP	58,77%	58,76%	58,76

4.5 Avaliação do desempenho dos algoritmos utilizando validação cruzada com janela expansiva

Nesta secção aplicou-se a técnica de validação cruzada com janela expansiva, permitindo avaliar o desempenho temporal dos algoritmos (fora da amostra) de cada par de moeda em múltiplos tamanhos de passos de expansão da amostra de treinamento (método explicito na secção 3.2.7.2).

EUR/USD

A Figura 4.18 ilustra a evolução das medidas obtidas via validação cruzada com janela expansiva à medida que a janela de treinamento se expande no algoritmo RL-L. Observa-se que a acurácia do algoritmo RL-L estático (59,95%, Tabela 4.2) permanece dentro do intervalo de confiança da validação (área sombreada), evidenciando a consistência temporal do algoritmo. A acurácia média da validação por tamanho de passo (k) situa-se, majoritariamente, acima de 0,60 (59,95%) obtido pelo RL-L estático. Tal comportamento indica que no retreinamento do algoritmo com as k observações subsequentes, o modelo mantém-se estável para dados fora da amostra, indicando robustez. Adicionalmente, o coeficiente de variação (CV) apresenta um decréscimo aproximadamente exponencial, estabilizando-se entre 0,04 e

0,10. Essa tendência reflecte menor sensibilidade do algoritmo à expansão da amostra de treino.

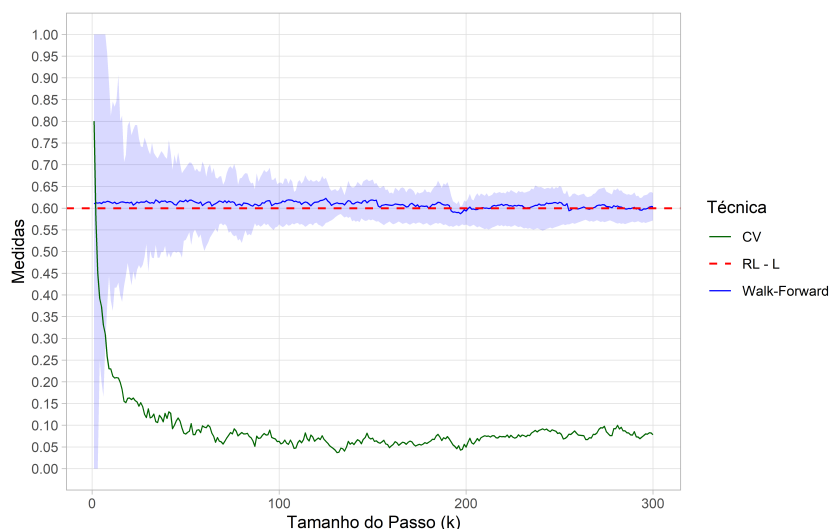


Figura 4.18: Desempenho da validação cruzada com janela expansiva no algoritmo RL-L para o par EUR/USD.

A Figura 4.19 ilustra os resultados da validação cruzada com janela expansiva no algoritmo de árvore de decisão (AD-C). Diferentemente do algoritmo anterior (Figura 4.18), observa-se que o desempenho médio da validação permanece consistentemente abaixo da acurácia observada pelo algoritmo AD-C estático (57,20%, Tabela 4.2). Verifica-se que após k passos a acurácia do algoritmo estático aparece fora, porém acima do intervalo de confiança. Quanto ao coeficiente de variação (CV), nota-se um decréscimo aproximadamente exponencial dos valores, com posterior estabilização entre 0 e 0,10. Essa estabilização reflecte a redução da variabilidade relativa e a consolidação da estabilidade predictiva do algoritmo à medida que a base de treinamento é ampliada.

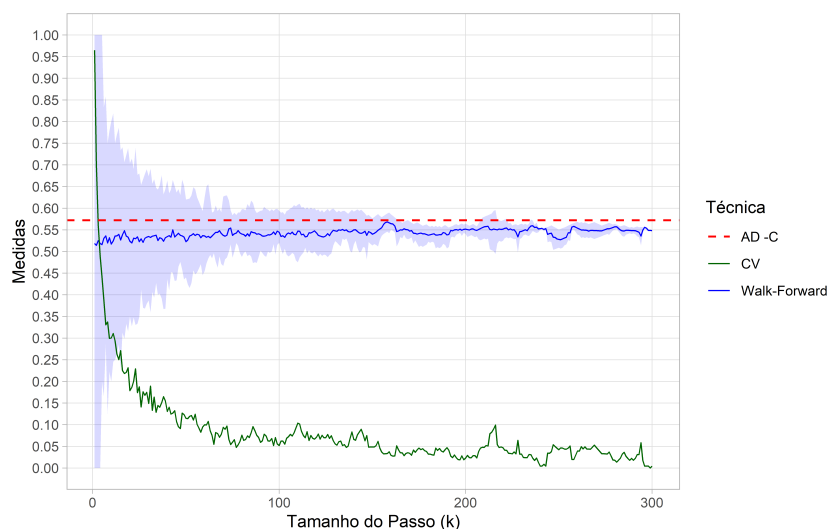


Figura 4.19: Desempenho da validação cruzada com janela expansiva no algoritmo AD-C para o par EUR/USD.

A Figura 4.20 ilustra a evolução das medidas obtidas pelo algoritmo RNA-BP através da validação cruzada com janela expansiva no par EUR/USD. Observa-se que a acurácia média da validação permanece ligeiramente abaixo do desempenho do modelo estático (61,39%, Tabela 4.2). Verifica-se também que, a acurácia do modelo estático situa-se sistematicamente acima do limite superior do intervalo de confiança da validação cruzada. Adicionalmente observa-se que o coeficiente de variação (CV) se estabiliza, neste cenário, entre 0 e 0,10, indicando menor sensibilidade do algoritmo à expansão da amostra de treino.

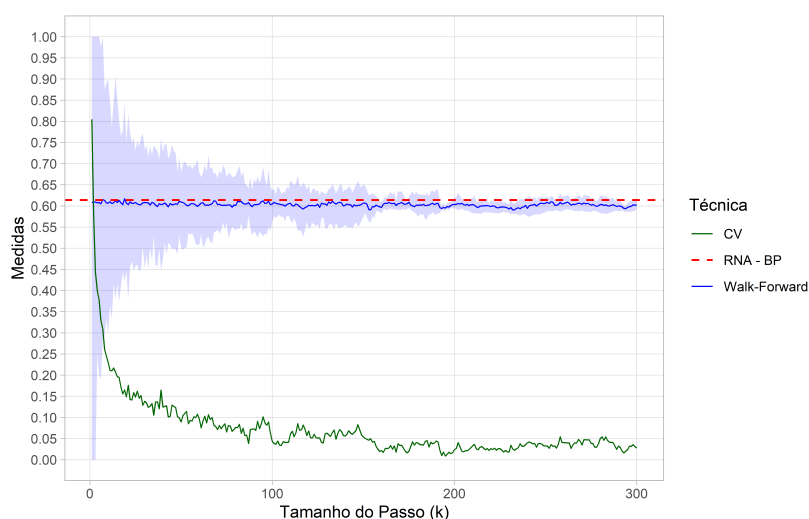


Figura 4.20: Desempenho da validação cruzada com janela expansiva no algoritmo RNA - BP para o par EUR/USD.

As Figuras 4.40, 4.41 e 4.42 (em anexo) apresentam a distribuição das acurácias médias obtidas via validação cruzada para o par EUR/USD. Os boxplots indicam que, os modelos apresentaram acurácia média em teste (fora da amostra) inferior à obtida em treino (dentro da amostra), com presença de valores extremos. O comportamento dessas distribuições, são detalhadas por meio de histogramas das referidas figuras.

GBP/USD

A Figura 4.21 apresenta a evolução das medidas do algoritmo RL-L sob validação cruzada com janela expansiva para o par GBP/USD. Observa-se que a acurácia média da validação permanece muito próxima ao longo dos diferentes tamanhos de passo (k), situando-se em torno da acurácia do algoritmo estático (61,26%, Tabela 4.2) e mantendo-se consistentemente dentro do intervalo de confiança estimado da validação. Esse comportamento evidencia a consistência temporal e predictiva fora da amostra à medida que a janela de treinamento é progressivamente ampliada.

O coeficiente de variação (CV) apresenta inicialmente um decréscimo acentuado, com posterior estabilização em níveis inferiores a 0,10. Tal dinâmica reflecte a redução progressiva da variabilidade à expansão da amostra de treino.

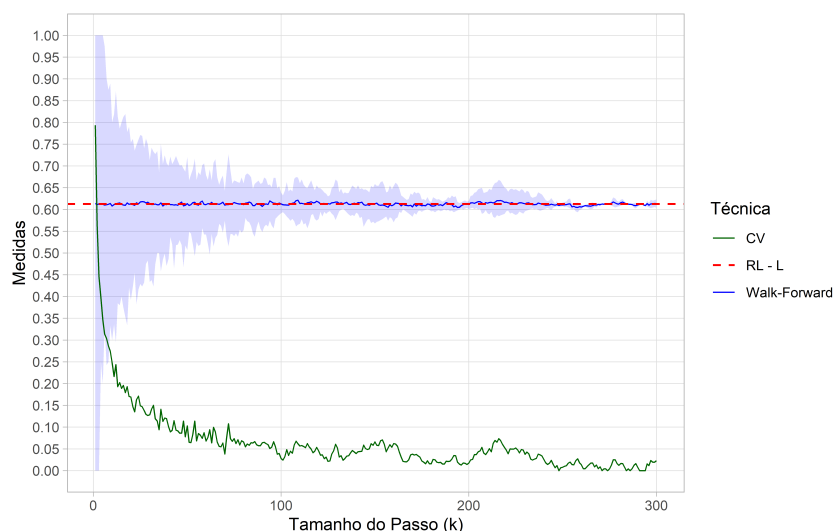


Figura 4.21: Desempenho da validação cruzada com janela expansiva no algoritmo RL-L para o par GBP/USD.

Na Figura 4.22 verifica-se que a validação para o algoritmo de árvore de decisão (AD-C) apresenta a acurácia média oscilando em torno da acurácia do algoritmo AD-C estático (54,7%, Tabela 4.2). Verifica-se que o algoritmo estático fica maioritariamente entre os intervalos de confiança. No que se refere ao coeficiente de variação, verifica-se um decréscimo aproximadamente exponencial dos valores, com posterior estabilização em níveis de 0 e 0,15.

Esse comportamento indica menor sensibilidade do algoritmo à expansão da amostra de treino.

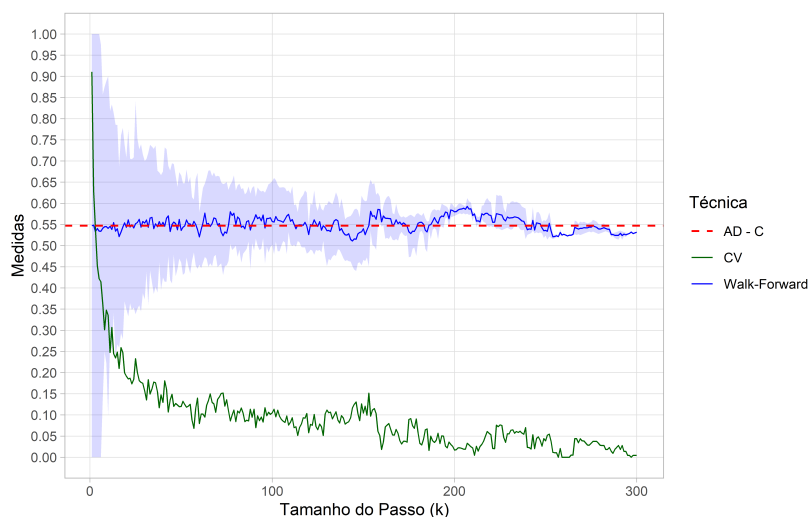


Figura 4.22: Desempenho da validação cruzada com janela expansiva no algoritmo AD-C para o par GBP/USD.

A Figura 4.23 ilustra a evolução das medidas obtidas pelo algoritmo RNA-BP através da validação cruzada com janela expansiva no par GBP/USD. Observa-se que a acurácia média da validação permanece acima do desempenho do modelo estático (60,08%, Tabela 4.2). Verifica-se que o algoritmo estático situa-se sistematicamente dentro do intervalo de confiança (95%) da validação. O coeficiente de variação oscila, neste cenário, entre 0 e 0,10, indicando menor sensibilidade do algoritmo à expansão da amostra de treino.

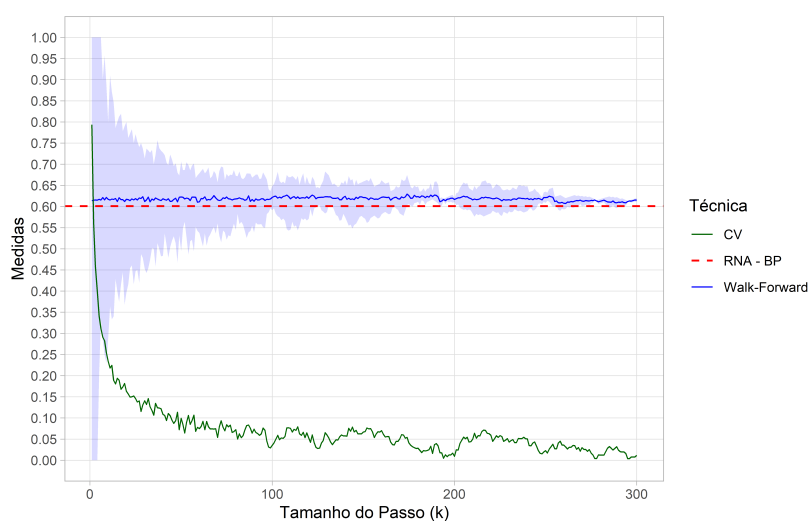


Figura 4.23: Desempenho da validação cruzada com janela expansiva no algoritmo RNA - BP para o par GBP/USD.

As Figuras 4.43, 4.44 e 4.45 (em anexo) apresentam a distribuição das acurácias médias obtidas via validação cruzada com janela expansiva para o par USD/JPY. Os boxplots indicam que, os modelos apresentaram acurácia média em teste (fora da amostra) inferior à obtida em treino (dentro da amostra), sem a presença de valores extremos. O comportamento dessas distribuições, são detalhadas por meio de histogramas, das referidas figuras.

USD/JPY

Na Figura 4.24, observa-se que a acurácia média da validação para o algoritmo RL-L no par USD/JPY permanece constantemente abaixo da acurácia do modelo estático (66,75%, Tabela 4.2). O algoritmo estático situa-se dentro do intervalo de confiança (95%) da validação em todo o domínio de k . O coeficiente de variação apresenta um decréscimo aproximadamente exponencial, estabilizando-se entre 0,05 e 0,10, o que indica menor sensibilidade do algoritmo à expansão da amostra de treino.

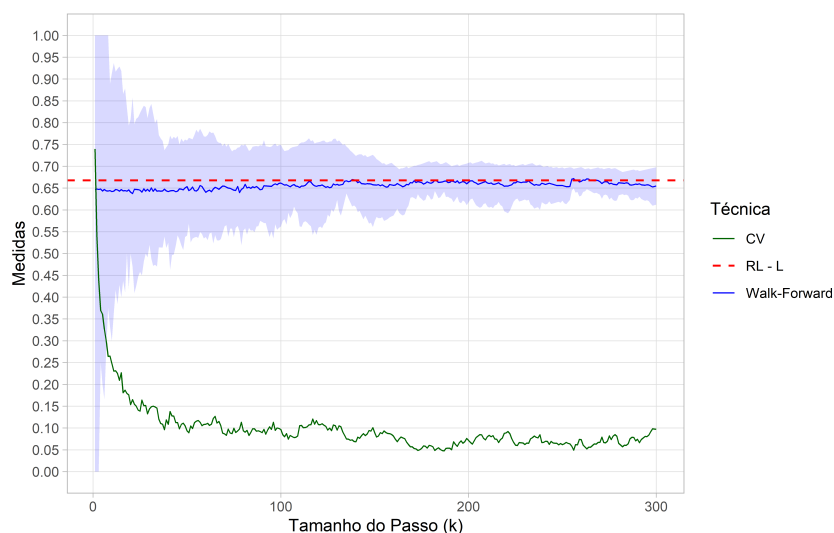


Figura 4.24: Desempenho da validação cruzada com janela expansiva no algoritmo RL-L para o par USD/JPY.

A Figura 4.25 apresenta os resultados da validação para o algoritmo de árvore de decisão (AD-C) aplicado ao par USD/JPY. Observa-se que a acurácia média da validação permanece sistematicamente inferior à do modelo estático (57,33%, Tabela 4.2), assim como o seu intervalo de confiança. O coeficiente de variação apresenta uma redução gradual, estabilizando-se entre 0 e 0,10 à medida que a janela de treino se expande. Este comportamento indica que o processo de retreino contribui para uma maior estabilidade das previsões, embora sem ganhos relevantes na acurácia média fora da amostra.

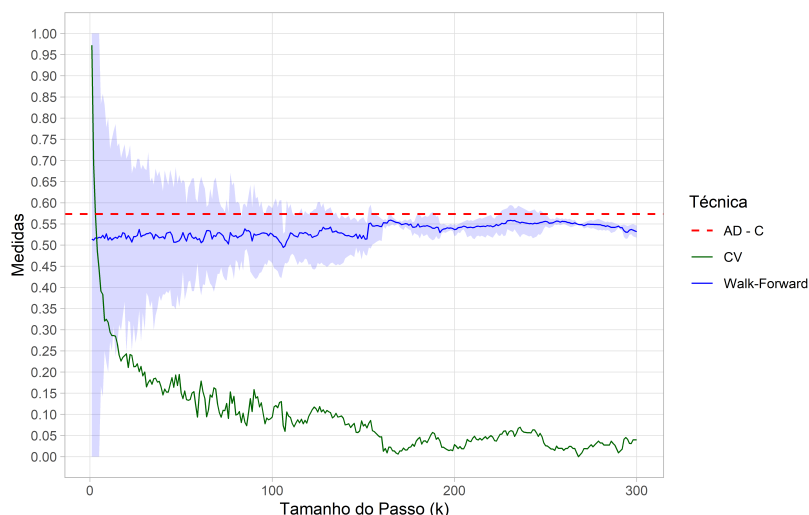


Figura 4.25: Desempenho da validação cruzada com janela expansiva no algoritmo AD-C para o par USD/JPY.

A Figura 4.26 ilustra a evolução das medidas obtidas pelo algoritmo RNA-BP através da validação cruzada com janela expansiva. Observa-se que a acurácia média da validação permanece sistematicamente abaixo do desempenho do modelo estático (61,78%, Tabela 4.2). Verifica-se também que, para $k > 150$, o intervalo de confiança da validação cruzada situa-se integralmente abaixo do valor da acurácia do modelo estático. O coeficiente de variação (CV) oscila, neste cenário, entre 0 e 0,15.

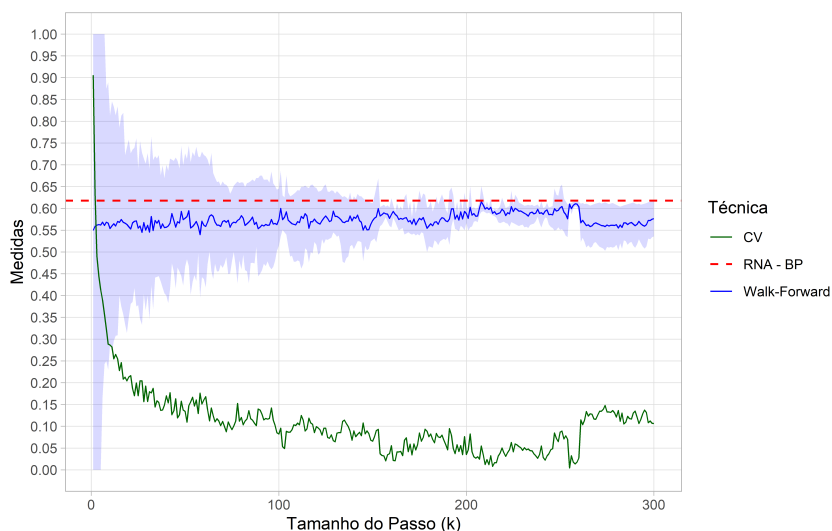


Figura 4.26: Desempenho da validação cruzada com janela expansiva no algoritmo RNA-BP para o par USD/JPY.

As Figuras 4.46, 4.47 e 4.48 (em anexo) apresentam a distribuição das acurácias médias obtidas via validação cruzada para o par USD/JPY. Os boxplots indicam que, à exceção

do algoritmo de regressão logística penalizada (Lasso), os modelos apresentaram acurácia média em teste (fora da amostra) inferior à obtida em treino (dentro da amostra), sem a presença de valores extremos. O comportamento dessas distribuições, são detalhadas por meio de histogramas, das referidas figuras.

4.6 Aplicação das estratégias de negociação e análise de risco financeiro

Nesta secção, avalia-se o desempenho financeiro das estratégias baseadas nos algoritmos de machine learning seleccionados. A selecção fundamentou-se nos critérios detalhados nas Tabelas 3.5 e 3.6. Para o par USD/JPY, o modelo escolhido foi a regressão logística penalizada via lasso (RL-L), uma vez que apresentou diferenças estatísticas significativas com o segundo melhor algoritmo (RNA - BP), comprovando sua superioridade em acurácia (4.3.4). Já para os pares EUR/USD e GBP/USD, seleccionaram-se as redes neurais artificiais (RNA). Justifica-se essa escolha pelo facto de os dois melhores algoritmos não apresentarem diferença estatística significativa nesses pares, nesses casos, o critério de desempate prioriza o modelo com maior retorno financeiro, conforme demonstrado nas figuras 4.58 e 4.59 em anexo.

As previsões dos modelos foram integradas às regras operacionais do Algoritmo 2 (detalhado na secção de Metodos), que combina os sinais gerados pelos modelos com filtros e estratégias de gestão de risco.

4.6.1 Resultados da estratégia de negociação no par GBP/USD

A Figura 4.27 apresenta a curva do retorno acumulado aritmético das negociações obtido pela aplicação da estratégia baseada nas previsões da rede neural artificial, em comparação ao desempenho do algoritmo sem a estratégia. Observa-se que a curva da estratégia proposta mantém-se sistematicamente superior a dos retornos acumulados da rede neural bruta.

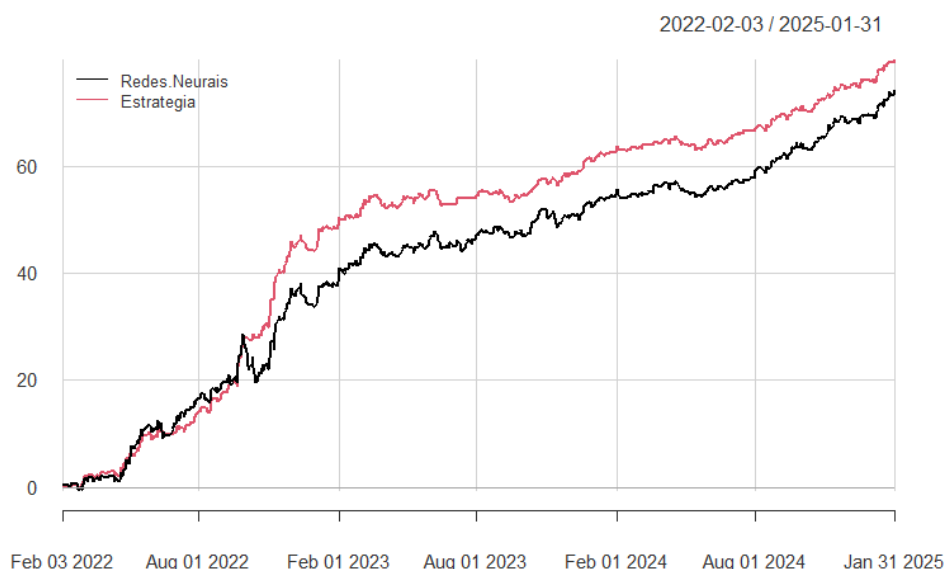


Figura 4.27: Comparação da curva do retorno diário acumulado das negociações com e sem a aplicação da estratégia baseada nas previsões da rede neural para o par GBP/USD.

Na Figura 4.28(a), os retornos anuais demonstram o desempenho da estratégia em relação ao sinal bruto da rede neural. Em 2022, o retorno total da estratégia foi de 44,18% contra 33,62% da rede neural bruta. Nos anos 2023, 2024 e 2025, a estratégia obteve um retorno total inferior ao do modelo bruto (16,66% contra 18,82%, 15,15% contra 16,97% e 3,96% contra 4,8%, respectivamente).

A Figura 4.28(b) detalha a frequência absoluta dos sinais operacionais. Foram identificados 342 sinais de compra, 258 de venda e 164 períodos de neutralidade. É relevante observar que, mesmo com 164 períodos sem exposição ao risco, a estratégia obteve um retorno acumulado superior ou aproximado, evidenciando a sua eficácia.

A Figura 4.28(c) ilustra o comportamento dos drawdowns, destacando que as perdas pontuais mais acentuadas ocorridas em 2022 e 2023 foram mitigadas pela estratégia.

Por fim, a Figura 4.28(d) apresenta a distribuição dos retornos, na qual se nota uma concentração de valores positivos e a presença de caudas longas favoráveis.

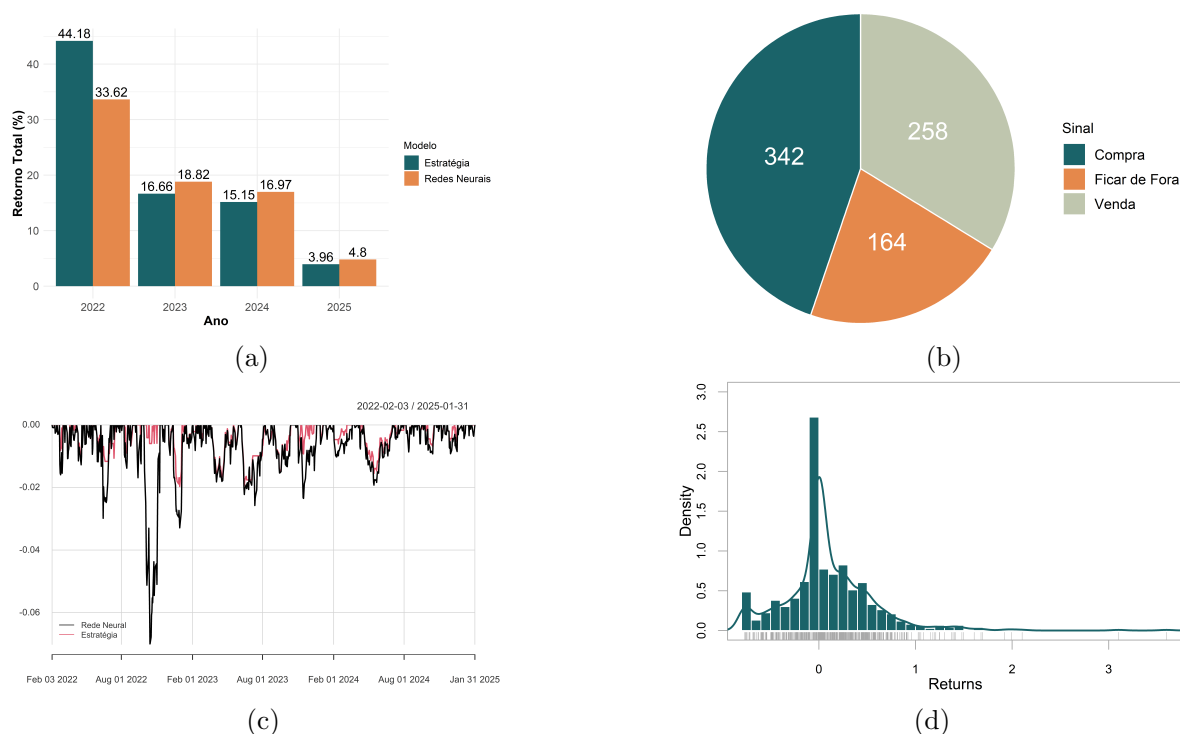


Figura 4.28: Desempenho da Estratégia no GBP/USD: (a) Retorno por ano; (b) Frequência absoluta dos sinais; (c) Drawdowns; (d) Distribuição dos retornos.

A Tabela 4.12 sintetiza as medidas de risco e desempenho da RNA-BP e da estratégia de negociação, evidenciando uma melhoria substancial no perfil de investimento desta última. Verifica-se que o retorno médio aumentou de 0,087% para 0,097% (aumento nominal de 0,01%), enquanto o desvio padrão dos retornos diminuiu de 0,576 para 0,468, respectivamente, implicando menor volatilidade e redução do risco. adicionalmente, a volatilidade anualizada da estratégia é de 7,433, reforçando o ganho em estabilidade e eficiência operacional, em relação ao sinal do algoritmo RNA-BP bruto. O drawdown máximo reduziu substancialmente, passando de -7,200% para -2,052%. Além disso, os índices de desempenho ajustados ao risco, Sharpe (2,075) e Sortino (4,103), superaram significativamente os obtidos pelos sinais brutos do RNA - BP (1,513 e 2,443, respectivamente), reforçando a eficiência da estratégia na otimização da relação risco-retorno.

Tabela 4.12: Medidas de desempenho da rede neural e da Estratégia no GBP/USD.

Medidas	RNA - BP	Estratégia
Drawdown Máximo	-7,200	-2,052
Retorno Médio	0,087	0,097
Curtose	6,722	10,209
Desvio Padrão dos Retornos	0,576	0,468
Desvio Padrão negativo	0,004	0,002
VaR 5%	-	-0,954
CVaR 5%	-	-1,498
Assimetria	0,345	1,462
Volatilidade Anualizada	9,140	7,433
Índice de Sharpe	1,513	2,075
Índice de Sortino	2,443	4,103
Retorno Acumulado	74,211	79,950

A assimetria evoluiu positivamente, de 0,345 para 1,462, indicando que a estratégia apresenta retornos positivos com maior frequência que os negativos. Esse é um factor favorável, pois demonstra a capacidade da estratégia em capturar movimentos de alta magnitude, mantendo as perdas controladas e limitadas. A curtose da estratégia foi de 10,209, reflectindo uma maior concentração de retornos em torno do retorno médio.

Em termos de medidas de risco de cauda, o VaR a 5% (utilizado como stop *loss fixo*) e o CVaR históricos situam-se em -0,954% e -1,498%, respectivamente. Assim, a estratégia mantém o risco potencial de perdas sob controle, mas com menor intensidade de drawdowns, conforme evidenciado pela redução do desvio padrão negativo de 0,004% para 0,002%.

O comportamento conjunto dessas medidas revela que a estratégia de negociação não apenas elevou o retorno acumulado (de 74,211% para 79,950%), como também optimizou o perfil de distribuição de risco, reduzindo a dispersão negativa e ampliando a probabilidade de ganhos assimétricos positivos.

4.6.2 Resultados da estratégia de negociação no par USD/JPY

A Figura 4.29 apresenta a curva do retorno acumulado aritmético das negociações no par USD/JPY, comparando o desempenho da estratégia e o sinal bruto do modelo. Verifica-se que a curva correspondente à estratégia apresenta um retorno superior ao longo do horizonte temporal.

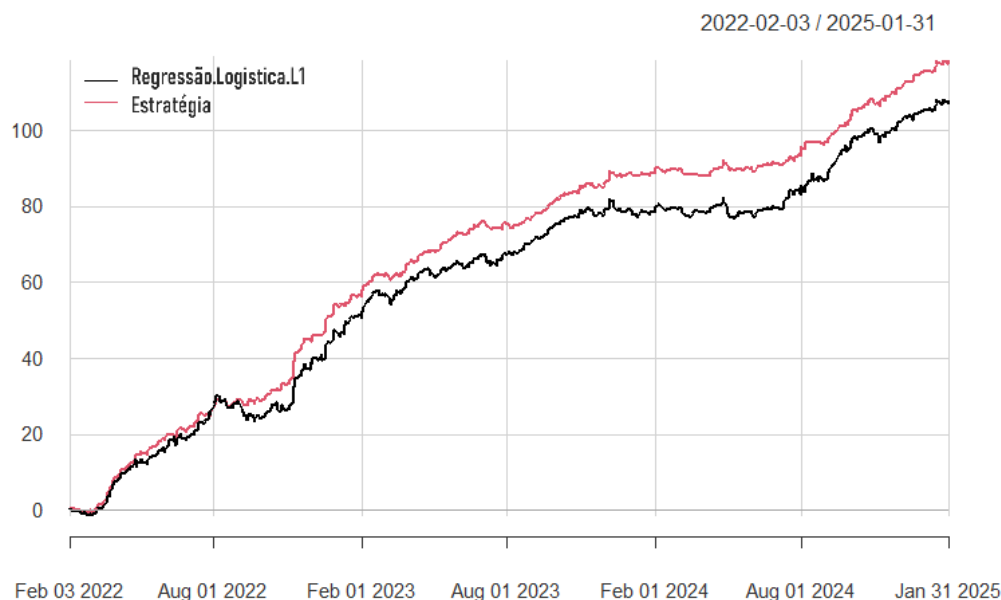


Figura 4.29: Comparação do retorno diário acumulado das negociações com e sem a aplicação da estratégia baseada nas previsões da regressão logística L1 para o par USD/JPY.

Na Figura 4.30(a), observa-se que a estratégia supera o modelo de regressão logística penalizada ($L1$) bruto em todo o período analisado. Em 2022, o retorno total atingiu 54,33% na estratégia contra 47,54% do sinal bruto do modelo. Em 2023, os retornos foram de 34,80% e 32,13%, respectivamente. A estratégia manteve a vantagem em 2024 (26,45% contra 25,67%) e no início de 2025 (2,72% contra 2,44%), confirmando maior consistência de desempenho.

A Figura 4.30(b) apresenta a frequência dos sinais operacionais, indicando 405 sinais de compra, 230 de venda e 129 sinais de neutralidade. Esses números indicam que em aproximadamente 17% do período a estratégia permaneceu fora do mercado, isto é, sem executar nenhuma negociação (trade).

Na Figura 4.30(c), observa-se que os drawdowns apresentam um comportamento mais amortecido na linha referente à estratégia. Isso demonstra que as perdas pontuais foram mitigadas pela estratégia, reduzindo a profundidade e a duração das quedas de capital.

Por fim, a Figura 4.30(d) apresenta a distribuição dos retornos, na qual se nota uma concentração de valores positivos de pequena e média magnitude, com uma cauda longa à direita. Tal característica indica a ocorrência ocasional de ganhos expressivos que contribuem para o retorno acumulado positivo final.

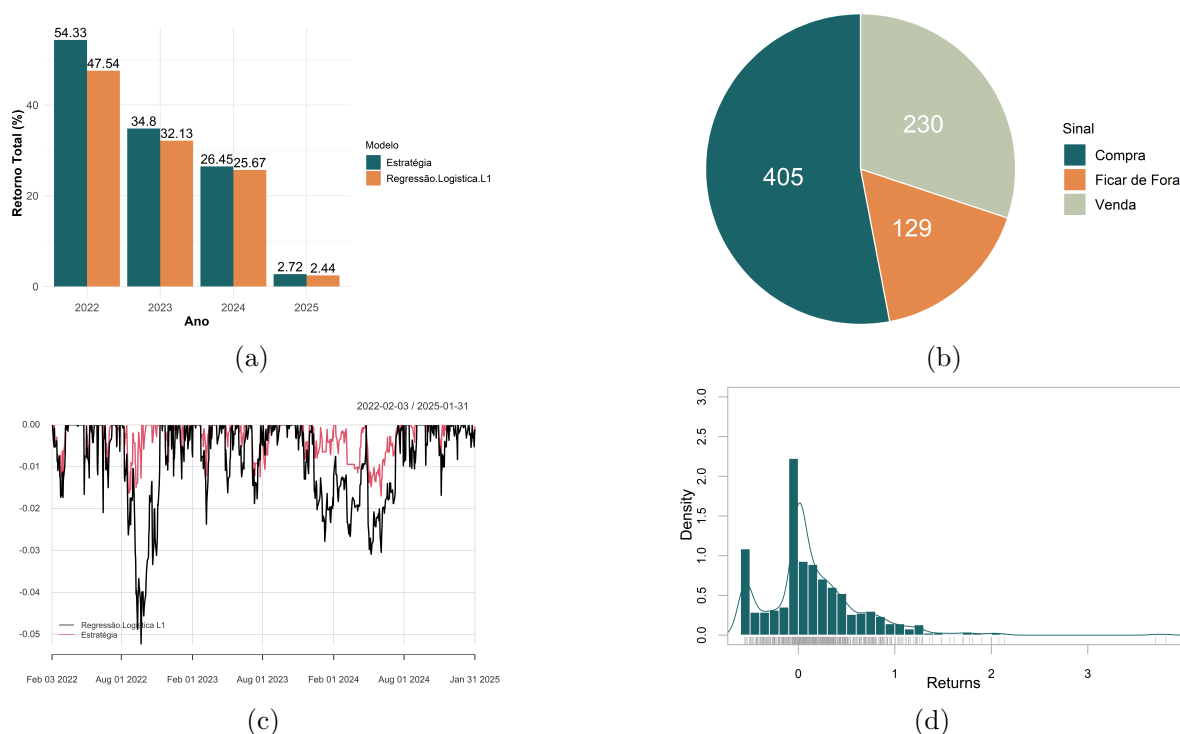


Figura 4.30: Desempenho da estratégia híbrida e/ou modelo puro no USD/JPY: (a) Retorno por ano; (b) Frequência de sinais; (c) Drawdowns; (d) Distribuição dos retornos.

A Tabela 4.13 sintetiza as medidas de risco e desempenho da regressão logística (L1) e da estratégia de negociação, evidenciando uma melhoria substancial no perfil de investimento desta última. O retorno médio aumentou de 0,141% para 0,155%, indicando um incremento no desempenho médio diário. Esse resultado é confirmado pelo retorno acumulado que saiu de 108,061% para 118,284%.

No que tange ao risco, observa-se uma redução significativa no drawdown máximo, que passou de -5,237% para -1,695%, refletindo uma melhoria relevante na preservação de capital. A volatilidade anualizada diminuiu de 10,37% para 8,046%, evidenciando menor variabilidade dos retornos. Da mesma forma o desvio padrão negativo apresentou queda acentuada (de 0,004% para 0,002%), reforçando a mitigação do risco de perdas de capital.

O VaR histórico e o CVaR mantiveram-se inalterados, sendo o VaR utilizado como parâmetro de stop loss. Quanto à distribuição dos retornos, a assimetria aumentou de 0,298 para 1,659 e a curtose de 6,120 para 10,169, indicando uma maior propensão a retornos positivos e uma distribuição leptocúrtica, com maior concentração de resultados próximos à média e caudas favoráveis.

Tabela 4.13: Medidas de desempenho da regressão logística L1 e da estratégia no par USD/JPY.

Métrica	RL - L	Estratégia
Drawdown máximo	-5,237	-1,695
Retorno medio	0,141%	0,155
Curtose	6,120	10,169
Desvio padrão dos retornos	0,653	0,507
Desvio padrão negativo	0,004	0,002
VaR (10%)	-	-2,216
CVaR (5%)	-	-1,933
Assimetria	0,298	1,659
Volatilidade anualizada	10,371	8,046
Índice de Sharpe	2,159	3,055
Índice de Sortino	3,635	7,252
Retorno Acumulado	108,061	118,284

Por fim, os índices ajustados ao risco demonstram a clara superioridade da estratégia em relação ao modelo bruto. O índice de Sharpe aumentou de 2,159 para 3,055, e o índice de Sortino praticamente dobrou (de 3,635 para 7,252). Tais resultados confirmam uma melhoria substancial na relação risco-retorno e validam a eficácia da estratégia na geração de retornos excedentes com menor exposição ao risco negativo.

4.6.3 Resultados da estratégia de negociação no par EUR/USD

A Figura 4.31 apresenta a curva do retorno acumulado das negociações para o par EUR/USD, comparando a estratégia, baseada na rede neural artificial (RNA -BP), ao desempenho do sinal bruto do modelo. Observa-se que a curva da estratégia supera consistentemente o modelo de rede neural bruta.

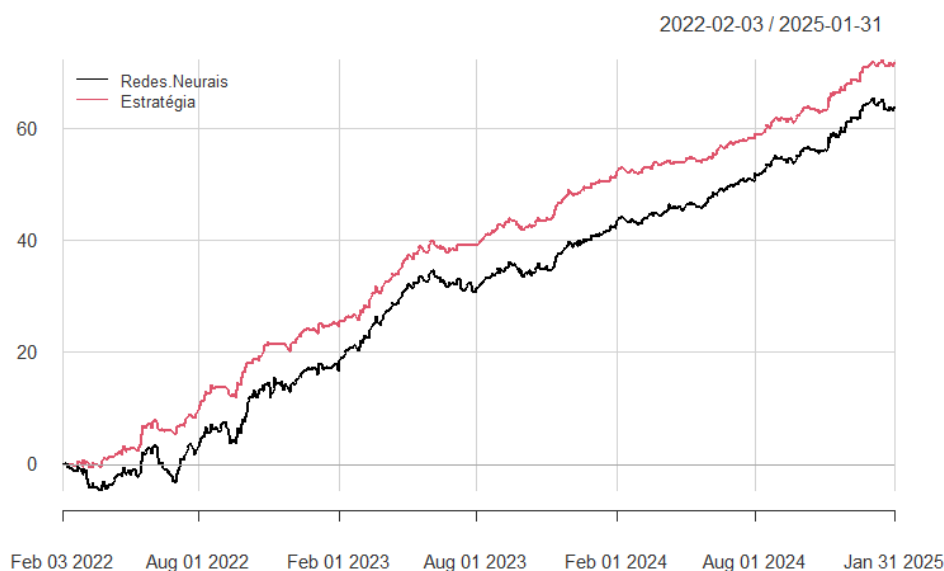


Figura 4.31: Comparação do retorno diário acumulado das negociações com e sem a aplicação da estratégia baseada nas previsões da rede neural para o par EUR/USD.

A Figura 4.32(a) apresenta os retornos anuais, destacando a superioridade da estratégia ao longo do período. Em 2022, o retorno total atingiu 24,26%, frente a 17,36% do modelo bruto. Em 2023, a estratégia apresentou retorno total de 25,91%, contra 23,46%. Contudo, em 2024, o sinal bruto da rede neural obteve um retorno relativamente superior (24,01%). No início de 2025, nota-se pequena variação negativa, a qual se mostrou mais controlada na estratégia.

O gráfico 4.32b representa a frequência absoluta dos sinais operacionais. Foram identificados 393 sinais de compra, 210 sinais de venda e 161 períodos de neutralidade. Estes períodos de neutralidade corroboram a superioridade da estratégia, demonstrando que, mesmo com 161 períodos sem operação, obteve-se um desempenho superior.

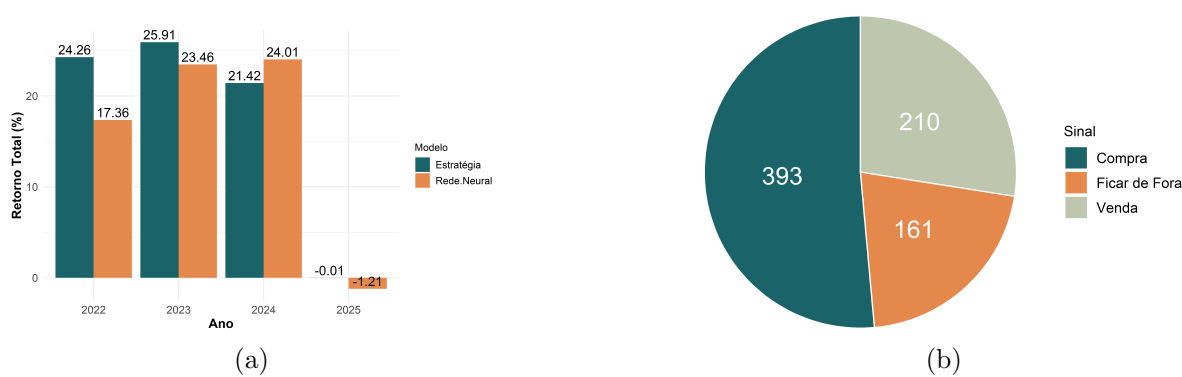


Figura 4.32: Desempenho da Estratégia no EUR/USD: (a) Retorno por ano; (b) Frequência absoluta dos sinais.

A Figura 4.33(a) ilustra o comportamento dos drawdowns, nos quais se nota que as quedas mais acentuadas no modelo base foram mitigadas pela estratégia. Já a Figura 4.33(b) exibe a distribuição dos retornos, com maior concentração de resultados positivos e cauda direita mais alongada, indicando que a estratégia foi eficaz em capturar ganhos de maior magnitude, enquanto limitou perdas extremas.

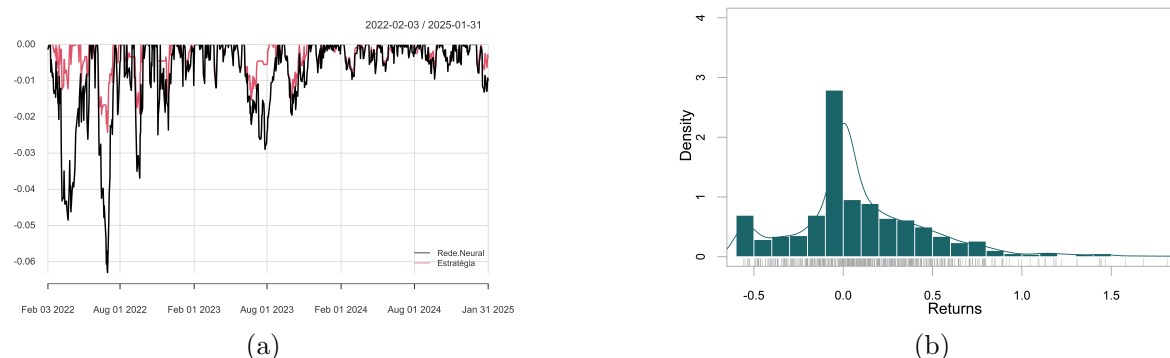


Figura 4.33: Desempenho da Estratégia no EUR/USD: (a) Drawdown; (b) Distribuição dos retornos.

Tabela 4.14: Medidas de desempenho da rede neural e da estratégia no par EUR/USD.

Métrica	RNA - BP	Estratégia
Drawdown Máximo	-6,3143	-2,424
Retorno Médio	0,083	0,094
Desvio Padrão dos Retornos	0,489	0,378
Desvio padrão negativo	0,003	0,002
VaR (5%)	-	-0,798
CVaR (5%)	-	-1,051
Assimetria	-0,138	0,873
Curtose	4,539	4,914
Volatilidade Anualizada	7,761	6,002
Índice Sharpe	1,703	2,478
Índice Sortino	2,670	4,985
Retorno Acumulado	63,621	71,580

Os resultados apresentados na Tabela 4.14 evidenciam ganhos significativos e eficiência operacional na aplicação da estratégia. O retorno médio aumentou de 0,083% para 0,094%, indicando um incremento no desempenho médio diário, pelo que se observa os efeitos no retorno acumulado (63,621% para 71,580%, respectivamente).

Quanto ao risco, observa-se uma redução expressiva no drawdown máximo da estratégia, que passou de 6,314% para 2,424%. Esse resultado reflecte uma maior preservação de capital, especialmente em períodos de maior volatilidade.

A volatilidade anualizada diminuiu de 7,761% para 6,002% na RNA-BP e na estratégia, respectivamente, reforçando o ganho de estabilidade desta última. O desvio padrão dos retornos apresentou queda de 0,489% para 0,378%, e o desvio padrão negativo reduziu-se de 0,003% para 0,002%, reforçando a mitigação do risco de perdas expressivas.

Quanto à distribuição dos retornos, a assimetria passou de -0,138 para 0,873, indicando uma maior propensão a retornos positivos. Já a curtose aumentou de 4,540 para 4,914, sugerindo uma distribuição mais concentrada em torno da média, com ocorrência de eventos extremos controlados, dado que o VaR (5%) fixo limitou o risco de perdas excessivas.

Por fim, os índices ajustados ao risco apresentaram melhorias expressivas, o índice de Sharpe passou de 1,703 para 2,478 e o índice de Sortino praticamente dobrou (de 2,670 para 4,985). Tais resultados confirmam uma melhoria substancial na relação risco-retorno e validam a eficácia da estratégia na geração de retornos excedentes com menor exposição ao risco de perda.

4.7 Discussão dos Resultados

As correlações observadas entre os pares EUR/USD e GBP/USD e o índice do dólar (DXY) reforçam o papel central do dólar norte-americano na dinâmica do mercado cambial global, em linha com evidências empíricas reportadas por Verdelhan (2018). É fundamental destacar que a robustez dessa correlação é matematicamente influenciada pela própria composição do índice, conforme aponta ICE (2024), o Euro detém o maior peso relativo no DXY (57,6%), seguido pela Libra Esterlina (11,9%), o que explica a sensibilidade do índice aos movimentos desses pares específicos.

Adicionalmente, a associação positiva entre USD/JPY e o DXY, deve-se ao facto de que o Iene compõe 13,6% do peso do índice DXY. Para além da dinâmica interna do índice, a correlação com o Nikkei 225 sugere uma interdependência entre o mercado de acções japonês e a taxa de câmbio local. Esse fenómeno é frequentemente atribuído a fluxos internacionais de capital e estratégias de *carry trade* (Burnside et al., 2011), onde a desvalorização do Iene tende a favorecer o sector exportador japonês, impulsionando o índice.

No que tange a análise de dois dos indicadores técnicos bem conhecidos, os resultados associados ao índice de força relativa evidenciam comportamentos assimétricos entre regimes de sobrecompra e sobrevenda. Observou-se que na maioria dos casos, níveis extremos de IFR tendem a estar associados a inclinações negativas nos retornos subsequentes em contextos de sobrecompra e inclinações positivas em situações de sobrevenda. Estes resultados estão alinhados com a interpretação clássica do IFR como indicador *contrarian* em mercados financeiros (Wilder, 1978; Achelis, 2001). Em contraste, o oscilador estocástico apresentou menor capacidade explicativa isolada, resultado coerente com estudos que apontam limitações desses indicadores quando utilizados de forma individual e sem combinação com outros sinais (Park & Irwin, 2007).

A análise de estacionaridade em janelas móveis revelou alternância entre períodos estacionários e não estacionários, sugerindo mudanças estruturais e regimes dinâmicos ao longo do tempo. Este comportamento é amplamente documentado na literatura sobre séries financeiras e está frequentemente associado a choques macroeconómicos e eventos extremos (Hamilton, 1989; Ang & Timmermann, 2012). De forma complementar, a evolução do coeficiente de Hurst indicou predominância de antipersistência em horizontes curtos, reforçando a evidência de reversão à média em determinados períodos.

A selecção das variáveis explicativas foi baseada em um critério empírico de redução de redundância, e não em testes de significância estatística individuais. Essa opção é consistente com a literatura que aponta limitações no uso de procedimentos de selecção baseados em p-valor em séries temporais financeiras, em função da dependência temporal, instabilidade

estrutural e da fraca associação entre significância estatística e capacidade predictiva (Granger, 1986; Freedman, 1991; López de Prado, 2020). Ressalta-se que este critério foi aplicado ao conjunto total de 209 variáveis analisadas neste trabalho.

Este estudo adopta uma abordagem predictiva de machine learning, na qual os algoritmos são utilizados como ferramentas de previsão e avaliados principalmente pelo desempenho fora da amostra, não tendo como objectivo a explicação estrutural das relações entre variáveis (Breiman, 2001). Em consonância com López de Prado (2018, 2020), no contexto específico dos mercados financeiros, os algoritmos predictivos devem ser apreciados sobretudo pela sua robustez empírica, desempenho fora da amostra e utilidade económica, em virtude da instabilidade estrutural dos dados financeiros (os detalhes dos modelos são apresentados no Apêndice A - Fig. 4.55, 4.56, 4.57 e B - Tabela 4.15).

No âmbito da modelagem predictiva, os algoritmos otimizados via grid search em conjunto com a validação cruzada demonstraram desempenho superior aos modelos sem optimização. A literatura corrobora que a optimização melhora o desempenho predictivo ao encontrar o equilíbrio ideal entre viés e variância (Bergstra, Bengio, 2012). No presente estudo, a eficácia deste processo foi robustecida pelo uso da validação cruzada para séries temporais, abordagem para garantir a dependência dos temporal em dados de series temporais (Packt, 2019).

A regressão logística penalizada via $L1$ destacou-se pela maior consistência e estabilidade fora da amostra. Para os pares USD/JPY e GBP/USD, a regressão logística $L1$ apresentou desempenho superior em termos de acurácia, superando os demais modelos. As redes neurais apresentaram desempenhos competitivos em alguns pares (GBP/USD e EUR/USD) em termos de acurácia, o que se verificou também nas medidas complementares (recall, precisão, especificidade e f1-score). Já a árvore de decisão apresentou o menor desempenho em todos os pares.

A robustez dos modelos foi corroborada tanto pelo bootstrap móvel por blocos quanto pela validação cruzada com janela expansiva. Em particular na validação cruzada com janela expansiva, observou-se que o aumento progressivo da amostra de treino contribuiu para a redução da variabilidade das previsões, com ganhos de acurácia média em alguns casos (EUR/USD e GBP/USD), sugerindo estabilidade da capacidade predictiva ao longo do tempo.

Os resultados do bootstrap móvel por blocos indicaram que o desempenho observado não é fruto de flutuações amostrais específicas, mas reflecte uma capacidade predictiva dos algoritmos relativamente estável sob dependência temporal.

A literatura financeira sugere que a utilidade prática de modelos predictivos depende menos da acurácia estatística isolada e mais da sua integração em estratégias de negociação com regras explícitas de decisão e gestão de risco. Conforme destaca Pardo (2011), um modelo de previsão, por mais acurado que seja, não constitui uma estratégia de investimento por si só, exigindo regras de execução e filtros que convertam sinais em operações de mercado.

No mercado de Forex, caracterizado por elevada liquidez e baixa previsibilidade direccional, estudos indicam que a optimização dos retornos está mais associada ao desenho da estratégia do que à complexidade dos modelos de previsão (Levich, 1998; Neely et al., 2009). Nesse contexto, a combinação de sinais predictivos com filtros operacionais e critérios de risco ajustado tem demonstrado maior capacidade de converter vantagens estatísticas modestas em desempenho económico consistente (Lo, 2004; López de Prado, 2018).

À luz desse enquadramento teórico, os resultados empíricos deste estudo mostraram que previsões estatisticamente moderadas podem ser economicamente ainda mais exploráveis quando combinadas com regras explícitas de gestão de risco, filtros e custos operacionais (estratégias de negociação). Observou-se que, apesar da redução do número de operações executadas, as estratégias apresentaram retorno acumulado superior ao obtido pelos algoritmos aplicados de forma directa, sem mecanismos de decisão adicionais. A redução consistente dos drawdowns, aliada à melhoria dos rácios de Sharpe e Sortino e à maior assimetria favorável da distribuição dos retornos, evidencia que a estratégia contribuiu para a optimização da relação risco-retorno. Esses resultados reforçam a noção de que, em mercados financeiros, pequenas vantagens predictivas podem ser convertidas em desempenho económico relevante quando integradas a estratégias disciplinadas de controlo de risco, em consonância com Lo (2004) & Bailey et al. (2014).

A incorporação de custos de transação foi fundamental para uma avaliação realista do desempenho das estratégias. De acordo com a literatura sobre, os custos de transação em pares altamente líquidos como EUR/USD, GBP/USD e USD/JPY, situam-se tipicamente entre 0.009% e 0.016% por operação via CFDs (Aliber et al., 2000; Lyons, 2001; Chelley-Steeley e Tsorakidis, 2013).

Em síntese, os resultados deste estudo indicam que a aplicação de técnicas de machine learning ao mercado Forex pode contribuir para a optimização, isto é, maximização dos retornos financeiros de investimentos quando combinada com procedimentos rigorosos de validação fora da amostra e estratégias de negociação estruturadas. Embora a previsibilidade direccional se mostre limitada, dada a Hipótese de Mercados Eficientes e a inerente natureza de passeio aleatório dos preços (Fama, 1970), os resultados empíricos sugerem que abordagens estatisticamente robustas permitem explorar essa previsibilidade de forma cautelosa e consistente.

Capítulo 5

CONCLUSÕES E RECOMENDAÇÕES

Este estudo teve como objectivo otimizar os retornos financeiros de investimentos no mercado de Forex por meio da aplicação de técnicas de machine learning. Os dados evidenciaram características típicas de séries financeiras, como elevadas mudanças de regime e ausência de estacionaridade persistente, confirmando a complexidade do mercado de Forex e justificando a utilização de abordagens predictivas flexíveis e métodos rigorosos de avaliação.

No que concerne à selecção do algoritmo que melhor se ajusta aos dados históricos, a regressão logística penalizada via Lasso obteve o melhor desempenho para o par USD/JPY. Por outro lado, para os pares EUR/USD e GBP/USD, as redes neurais artificiais destacaram-se como os modelos mais eficazes.

Por fim, a incorporação das previsões dos modelos em uma estratégia de negociação, fundamentada em regras explícitas de gestão de risco, constituiu uma metodologia que permitiu otimizar a relação risco-retorno das operações. Os resultados demonstram que, mesmo com acurácias moderadas, é possível obter ganhos económicos relevantes quando as estimativas são utilizadas como suporte à tomada de decisão.

Conclui-se que os algoritmos de machine learning podem aprimorar a análise e o processo decisório no mercado de Forex até um certo limite, contribuindo para a maximização dos retornos financeiros de forma incremental. Contudo, tais ferramentas não eliminam a incerteza e o carácter estocástico inerentes a esse mercado.

5.1 Recomendações

Como recomendações para trabalhos futuros, sugere-se:

- A exploração de horizontes de previsão multiperíodo, avaliando a consistência da acurácia e o comportamento dos algoritmos ao tentar prever a direcção dos retornos em janelas temporais mais longas;
- A incorporação de variáveis macroeconómicas e indicadores de regime de mercado, com o objectivo de capturar mudanças estruturais e dinâmicas temporais nos mercados cambiais;
- O estudo de abordagens explícitas de determinação do tamanho de cada operação individual, isto é, do nível de capital ou de exposição assumida em cada operação, com o objectivo de avaliar o impacto do dimensionamento das posições sobre a estabilidade e o desempenho económico das estratégias propostas.

REFERÊNCIAS

- Achelis, S. B. (2001). *Technical Analysis from A to Z* (2nd ed.). McGraw-Hill.
- Adariani, M. R. (2012). *Evaluation of the profitability of technical analysis for Asian currencies in the Forex spot market for short-term trading*. *AU-GSB e-JOURNAL*, 5(4), 4–18.
- Adnan, M., Khan, M. A., e Rauf, A. (2023). *Testing the Random Walk Hypothesis in Pakistan's Forex Market: An analysis of major currency dynamics*. *Journal of Economic Policy Research*, 10(2), 55–72.
- Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Wiley-Interscience.
- Agyei, S. K., Sun, S., Abor, J. Y., e Donkor, J. (2023). *Revisiting the efficient market hypothesis in African foreign exchange markets*. *Journal of International Financial Markets, Institutions and Money*, 84, 101731.
- Aliber, R. Z., Chowdhry, B., e Yan, S. (2000). *Some tests of purchasing power parity*. *The Journal of International Money and Finance*, 19(5), 725–737.
- Anastasakis, L., e Mort, N. (2009). *Exchange rate forecasting using a combined parametric and nonparametric self-organising modelling approach*. *Expert Systems with Applications*, 36(12), 12001–12011.
- Anderson, R. (2007). *The credit scoring toolkit: Theory and practice for retail credit risk management and decision automation*. Oxford University Press.
- Ang, A., e Timmermann, A. (2012). *Regime changes and financial markets*. *Annual Review of Financial Economics*, 4(1), 313–337.
- Archer, M. D. (2010). *Getting Started in Currency Trading*. Wiley.
- Arlot, S., e Celisse, A. (2010). *A survey of cross-validation procedures for model selection*. *Statistics Surveys*, 4, 40-79.
- Assumpção Júnior, C. B. (2009). *Um estudo do impacto da governança corporativa na rentabilidade*.

bilidade e performance do índice Brasil (IBrX). Trabalho de Conclusão de Curso, Fundação Instituto Capixaba de Pesquisa em Contabilidade, Economia e Finanças.

Atiya, A. F., e Magdon-Ismael, M. (2018). *The maximum drawdown of discrete time process. Advanced Mathematical Models and Applications*, 3(2), 95-105.

Atkinson, A. C. (2021). *Transformations and Outliers in Statistical Modeling*. London School of Economics.

Ayub, U., Ahsan, S. M., e Qureshi, S. M. (2022). Scalable big data pipeline for video stream analytics over commodity hardware. *KSII Transactions on Internet and Information Systems (TIIS)*, 16(4), 1146–1165.

Baasher, A. A., e Fakhr, M. W. (2010). *Forex Trend Classification using Machine Learning Techniques*.

BabyPips. (2015). *Learn Forex trading at the School of Pipsology*. <https://www.babypips.com/school/>. Acessado em 05 de novembro de 2025.

Bailey, D. H., Borwein, J. M., Lopez de Prado, M., e Zhu, Q. J. (2014). *Pseudo-mathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance. Notices of the American Mathematical Society*, 61(5), 458–471.

Bank for International Settlements. (2013). *Foreign exchange turnover in April 2013: Preliminary global results (Technical report)*. Bank for International Settlements.

Bank for International Settlements. (2025). *OTC foreign exchange turnover in April 2025*. BIS.

Barclaycobertura. (2018). *Institutional investors' adoption of algorithmic trading and quantitative strategies*.

Bergstra, J., e Bengio, Y. (2012). *Random search for hyper-parameter optimization. Journal of Machine Learning Research*, 13(2), 281–305.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Blanchard, O. (2017). *Macroeconomia*. Pearson.

Blau, W. (2001). *Momentum, direction, and divergence: 21st century models and strategies for managing options and futures*. Wiley.

Bodie, Z., Kane, A., e Marcus, A. J. (2014). *Investimentos (10ª ed.)*. AMGH.

Bollinger, J. (2001). *Bollinger on Bollinger Bands*. McGraw-Hill.

- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., e Ljung, G. M. (2015). *Time series analysis: Forecasting and control (5th ed.)*. Wiley.
- Bredin, D., e Hyde, S. (2004). *Forex risk: Measurement and evaluation using value-at-risk*. *Journal of Business Finance e Accounting*, 31(9-10), 1389–1417.
- Breiman, L., Friedman, J. H., Olshen, R. A., e Stone, C. J. (1984). *Classification and regression trees*. Wadsworth International Group.
- Burnside, C., Eichenbaum, M., e Rebelo, S. (2011). *Carry trade and momentum in currency markets*. *Annual Review of Financial Economics*, 3(1), 511–535.
- Cai, J., e Tan, K. S. (2007). *Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures*. *ASTIN Bulletin*, 37(1), 93–112.
- Caporale, G. M., Hunter, J., e Menla Ali, F. (2014). *On the linkages between stock prices and exchange rates: Evidence from the banking crisis of 2007–2010*. *International Review of Financial Analysis*, 33, 87–103.
- Cartridge, J., e Jessen, A. (2011). *CFDs made simple: A beginner's guide to Contracts for Difference success*. John Wiley e Sons.
- Cerqueira, V., Torgo, L., e Mozetič, I. (2019). *Evaluating time series forecasting models: An empirical study on performance estimation methods*. *Machine Learning*, 108(11), 1997–2028.
- Chaikin, M. (1992). *Chaikin Money Flow*. In *Technical Analysis of Stocks and Commodities Magazine*.
- Chande, T. S., e Kroll, S. (1994). *The new technical trader: Boost your profits by plugging into the latest indicators*. Wiley.
- Chatfield, C. (2003). *The analysis of time series: An introduction (6th ed.)*. Chapman e Hall/CRC.
- Chatfield, C. (2004). *The analysis of time series: An introduction (6th ed.)*. Chapman e Hall/CRC.
- Chekhlov, A., Uryasev, S., e Zabarankin, M. (2005). *Drawdown measure in portfolio optimization*. *International Journal of Theoretical and Applied Finance*, 8(1), 13–58.
- Chelley-Steeley, P. L., e Tsorakidis, N. (2013). *Herding, asymmetric information and the bid-ask spread in the foreign exchange market*. *Journal of International Financial Markets, Institutions and Money*, 24, 83–102.

- Chen, S.-H., Liao, C.-C., Raymond, C. P., e Wang, S.-T. (2019). *Fundamental analysis and its effectiveness in financial markets*. *Journal of Financial Studies*, 27(2), 1-25.
- Chen, Y., e Rogoff, K. (2003). *Commodity currencies*. *Journal of International Economics*, 60(1), 133–160.
- Chihab, Y., Bousbaa, Z., Chihab, M., Bencharef, O., e Ziti, S. (2019). Algo-trading strategy for intraweek foreign exchange speculation based on random forest and probit regression. *Applied Computational Intelligence and Soft Computing*, 2019.
- Conover, W. J. (1999). *Practical nonparametric statistics (3rd ed.)*. Wiley.
- Cotter, J., e Dowd, K. (2011). *Extreme spectral risk measures: An application to futures clearinghouse margin requirements*.
- Cuomo, S., Di Somma, V., e Piccialli, F. (2018). *Pricing estimation of a barrier option in an IoT scenario*. *Future Generation Computer Systems*, 110, 407–412.
- Davison, A. C., e Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press.
- DeLong, E. R., DeLong, D. M., e Clarke-Pearson, D. L. (1988). *Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach*. *Biometrics*, 44(3), 837–845.
- Dickey, D. A., e Fuller, W. A. (1979). *Distribution of the estimators for autoregressive time series with a unit root*. *Journal of the American Statistical Association*, 74(366), 427–431.
- Dolan, B. (2011). *Currency Trading for Dummies*. Wiley.
- Dougherty, J., Kohavi, R., e Sahami, M. (1995). *Supervised and unsupervised discretization of continuous features*. ICML.
- Dowd, K. (2002). *Measuring market risk*. Wiley.
- Duda, R. O., Hart, P. E., e Stork, D. G. (2001). *Pattern classification (2nd ed.)*. Wiley-Interscience.
- Efron, B. (1979). *Bootstrap methods: Another look at the jackknife*. *The Annals of Statistics*, 7(1), 1–26.
- Enders, W. (2014). *Applied Econometric Time Series*. 4th ed. Hoboken, NJ: Wiley.
- Engel, C., Mark, N. C., e West, K. D. (2015). *Factor model forecasts of exchange rates*. *Econometric Reviews*, 34(1–2), 32–55.

- Enkhbayar, S., e Ślepaczuk, R. (2025). Predictive modeling of foreign exchange trading signals using machine learning techniques. *Expert Systems with Applications*, 259, 124968.
- Fabozzi, F. J. (2007). *Fixed Income Analysis*. 2nd ed. Hoboken, NJ: John Wiley e Sons.
- Fabozzi, F., Focardi, S., e Kolm, P. (2014). *Quantitative equity investing: Techniques and strategies*. Wiley.
- Faghihi, F., Khalafi, H., e Mirvakili, S. M. (2011). *A literature survey of neutronics and thermal-hydraulics codes for investigating reactor core parameters: Artificial neural networks as the VVER-1000 core predictor*. In Nuclear Power Systems Simulations and Operations (pp. 103–122). InTech.
- Fama, E. F. (1970). *Efficient capital markets: A review of theory and empirical work*. *The Journal of Finance*, 25(2), 383–417.
- Fama, E. F., e French, K. R. (1988). *Permanent and temporary components of stock prices*. *Journal of Political Economy*, 96(2), 246–273.
- Fama, E. F., e French, K. R. (1996). *Multifactor explanations of asset pricing anomalies*. *Journal of Finance*, 51(1), 55–84.
- Farinha Rodrigues, J. P. (2015). *Medição e optimização do risco em selecção de carteiras (Tese de Mestrado)*. Universidade de Coimbra.
- Fatum, R., e Yamamoto, Y. (2016). *Intra-day signatures of currency interventions: The case of Japan*. *Journal of Financial Markets*, 29, 46–70.
- Fawcett, T. (2006). *An introduction to ROC analysis*. *Pattern Recognition Letters*, 27(8), 861–874.
- Fischer, T., e Krauss, C. (2018). *Deep learning for stock selection*. *European Journal of Operational Research*, 270(1), 396–408.
- Freedman, D. A. (1991). *Statistical models and shoe leather*. *Sociological Methodology*, 21, 291–313.
- Garcia, M. T. M., e Rodrigues, A. C. G. (2019). *The dynamic relationship between stock market indexes and foreign exchange (Working Paper REM 2019/90)*. ISEG – Lisbon School of Economics and Management, REM, Universidade de Lisboa.
- Geronazzo, T. R. (2019). *Análise de estratégias de investimento e medidas de risco no mercado financeiro [Dissertação de mestrado, Universidade de São Paulo]*.

- Gilda, S., Heidrich, B., e Kiraly, F. (2024). *tsbootstrap: Enhancing Time Series Analysis with Advanced Bootstrapping Techniques*. arXiv.
- Gitman, L. J., e Joehnk, M. D. (2005). *Princípios de investimentos (8ª ed.)*. Pearson.
- Goldberg, L. R., e Mahmoud, O. (2016). *Drawdown: From practice to theory and back again*. *Mathematics and Financial Economics*, 10(3), 223–257.
- Good, I. J. (1950). *Probability and the weighing of evidence*. Charles Griffin.
- Goodfellow, I., Bengio, Y., e Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.
- Graham, B. (2006). *The intelligent investor: The definitive book on value investing*. HarperBusiness.
- Granger, C. W. J. (1986). *Developments in the study of cointegrated economic variables*. *Oxford Bulletin of Economics and Statistics*, 48(3), 213–228.
- Gu, S., Kelly, B., e Xiu, D. (2020). *Empirical asset pricing via machine learning*. *The Review of Financial Studies*, 33(5), 2223–2273.
- Guida, T. (2018). *Big data and machine learning in quantitative investment*. Wiley.
- Guo, W., Zhang, D., e Zhang, X. (2023). *Risk modeling and tail-event prediction using quantile-based approaches*. *Quantitative Finance*, 23(4), 721–738.
- Gupta, A., e Mishra, R. (2024). *A comprehensive study of stock market volatility: Types, determinants, and measurement methods*. *International Journal for Multidisciplinary Research (IJFMR)*, 5(2), 1-10.
- Guyard, K. C., e Deriaz, M. (2024). Predicting Foreign Exchange EUR/USD Direction Using Machine Learning. *Proceedings of the 2024 7th International Conference on Machine Learning and Machine Intelligence (MLMI)*, 1–9.
- Hajizadeh, E., Mahootchi, M., Esfahanipour, A., e Kh, M. M. (2019). *A new NN-PSO hybrid model for forecasting Euro/Dollar exchange rate volatility*. *Neural Comput. Appl.*, 31, 2063–2071.
- Hamilton, J. D. (1989). *A new approach to the economic analysis of nonstationary time series and the business cycle*. *Econometrica*, 57(2), 357–384.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.
- Han, J., Kamber, M., e Pei, J. (2012). *Data Mining: Concepts and Techniques*. Morgan

Kaufmann.

Hastie, T., Tibshirani, R., e Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.

Huang, R. D., e Masulis, R. W. (1999). *FX spreads and dealer competition across the 24-h trading day*. *Review of Financial Studies*, 12(1), 61–93.

Hull, J. C. (2015). *Options, futures, and other derivatives* (9th ed.). Pearson.

Hull, J. C. (2018). *Options, futures, and other derivatives* (10th ed.). Pearson.

Hyndman, R. J., e Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts.

Hyndman, R. J., e Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). Melbourne: OTexts.

Ince, H., e Trafalis, T. B. (2006). *A hybrid model for exchange rate prediction*. *Decis. Support Syst.*, 42, 1054–1062.

Intercontinental Exchange. (2024). *U.S. Dollar Index (USDX) methodology*.

Izbicki, R., e Santos, T. M. (2018). *Aprendizado de máquina: Uma abordagem estatística*. Edição dos autores.

James, G., Witten, D., Hastie, T., e Tibshirani, R. (2014). *An introduction to statistical learning: with applications in R*. Springer.

Jensen, M. C. (1978). *Some anomalous evidence regarding market efficiency*. *Journal of Financial Economics*, 6(2), 95–101.

Jondeau, E., Poon, S.-H., e Rockinger, M. (2007). *Financial modeling under non-gaussian distributions*. Springer.

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3^a ed.). McGraw-Hill.

Jubert de Almeida, B., Ferreira Neves, R., e Horta, N. (2018). Combining support vector machine with genetic algorithms to optimize investments in Forex markets with high leverage. *Applied Soft Computing*, 64, 596–613.

Kira, K., e Rendell, L. A. (1992). *A practical approach to feature selection*. In *Proceedings of the Ninth International Workshop on Machine Learning* (pp. 249–256).

- Kisiala, J. (2015). *Conditional Value-at-Risk: Theory and applications [Master's thesis, University of Edinburgh]*.
- Koller, D., e Sahami, M. (1996). *Toward optimal feature selection*. In Proceedings of the International Conference on Machine Learning.
- Kondratenko, V. Y., e Kuperin, Y. A. (2003). *Self-organization of the foreign exchange market: A dynamical model of the distribution of agents*. *Journal of Physics A: Mathematical and General*, 36(43), 10839–10851.
- Kotlan, Joanes, D. N., e Gill, C. A. (1998). *Comparing measures of sample skewness and kurtosis*. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 183–189.
- Kotsiantis, S., e Kanellopoulos, D. (2006). *Discretization techniques: A recent survey*.
- Kritzer, A. (2012). *Forex for Beginners: A Comprehensive Guide to Profiting from the Global Currency Markets*. Apress.
- Kuhn, M., e Johnson, K. (2013). *Applied predictive modeling*. New York: Springer.
- Künsch, H. R. (1989). *The jackknife and the bootstrap for general stationary observations*. *The Annals of Statistics*, 17(3), 1217–1241.
- Lahiri, S. N. (2003). *Resampling methods for dependent data*. Springer.
- Lane, G. C. (1984). *Lane's stochastics*. *Technical Analysis of Stocks e Commodities*, 2(3), 87-90.
- Levich, R. M. (1998). *International Financial Markets: Prices and Policies*. Irwin/McGraw-Hill.
- Levinson, M. (2014). *The Economist guide to financial markets: Why they exist and how they work*. The Economist: London, UK.
- Liu, F. T., Ting, K. M., e Zhou, Z.-H. (2008). *Isolation Forest*. 2008 Eighth IEEE International Conference on Data Mining.
- Liu, H., Hussain, F., Tan, C., e Dash, M. (2002). *Discretization: An enabling technique*. Data Mining and Knowledge Discovery.
- Lo, A. W. (2002). *The Statistics of Sharpe Ratios*. *Financial Analysts Journal*, 58(4), 36–52.
- Lo, A. W. (2004). *The Adaptive Markets Hypothesis: Market efficiency from an evolutionary perspective*. *Journal of Portfolio Management*, 30(5), 15–29.

- Lo, A. W., e MacKinlay, A. C. (1988). *Stock market prices do not follow random walks: Evidence from a simple specification test*. *The Review of Financial Studies*, 1(1), 41–66.
- López de Prado, M. (2018). *Advances in financial machine learning*. Wiley.
- López de Prado, M. (2020). *Machine learning for asset managers*. Cambridge University Press.
- López-Herrera, F., Venegas-Martínez, F., e Cruz-Aké, S. (2025). Directional forecasting for eight Forex pairs against the US dollar using machine learning techniques. *Discover Artificial Intelligence*, 5(224).
- Lyons, R. K. (2001). *The microstructure approach to exchange rates*. MIT Press.
- Madura, J. (2020). *International Financial Management (13th ed.)*. Cengage Learning.
- Makridakis, S., Wheelwright, S. C., e Hyndman, R. J. (1998). *Forecasting: Methods and applications (3rd ed.)*. Wiley.
- Malkiel, B. G. (2003). *The efficient market hypothesis and its critics*. *Journal of Economic Perspectives*, 17(1), 59–82.
- Markowitz, H. (1952). *Portfolio selection*. *The Journal of Finance*, 7(1), 77–91.
- Masini, R., Medeiros, M. C., e Mendes, E. F. (2020) *Machine learning advances for time series forecasting*. European Journal of Operational Research.
- Masry, S., Dupuis, A., Olsen, R., e Tsang, E. (2013). *Time zone normalization of FX seasonality*. *Quantitative Finance*, 13(1), 1115–1123.
- McNemar, Q. (1947). *Note on the sampling error of the difference between correlated proportions or percentages*. *Psychometrika*, 12(2), 153–157.
- Menezes, A. (2011). *Investimentos: Teoria e prática*. Atlas.
- Menkhoff, L., Schmidt, U., e Zhou, X. Y. (2018). *Does technical analysis really work? Evidence from foreign exchange markets*. *Journal of Economic Literature*, 56(4), 1205–1231.
- MetaQuotes. (2021). *MetaQuotes is a leading developer of software applications for financial markets*. MetaTrader 5. Acessado em 2 de setembro de 2025. Disponível em: <https://www.metatrader5.com/en/company>.
- Mishkin, F. S. (2015). *The Economics of Money, Banking, and Financial Markets (10th ed.)*. Pearson.

- Mishkin, F. S. (2016). *The economics of money, banking, and financial markets (11th ed.)*. Pearson Education.
- Montgomery, D. C., e Runger, G. C. (2014). *Applied statistics and probability for engineers (6th ed.)*. Wiley.
- Morettin, P. A. (2011). *Estatística nas finanças: Uma introdução à econometria financeira (2ª ed.)*. Blucher.
- Morettin, P. A., e Toloi, C. F. C. (2006). *Análise de séries temporais (2ª ed.)*. Edgard Blücher.
- Murphy, J. J. (1991). *Intermarket technical analysis: Trading strategies for the global stock, bond, commodity, and currency markets*. Wiley.
- Murphy, J. J. (1999). *Technical Analysis of the Financial Markets*. New York Institute of Finance.
- Murphy, J. J. (2015). *Trading with intermarket analysis: A visual framework for strategies in stocks, currencies, commodities, and bonds*. Wiley.
- Narendra, P. M., e Fukunaga, K. (1977). *A branch and bound algorithm for feature selection*. IEEE Transactions on Computers, C-26(9), 917–922.
- Neely, C. J., Weller, P. A., e Ulrich, J. M. (2009). *The adaptive markets hypothesis: Evidence from the foreign exchange market*. *Journal of Financial and Quantitative Analysis*, 44(2), 467–488.
- Nguyen Thi Thu, T., e Xuan, V. D. (2018). Supervised support vector machine in predicting foreign exchange trading. *International Journal of Engineering & Technology*, 7(9), 400-404.
- Norman, D. J. (2009). *CFDs: The definitive guide to Contracts for Difference*. Harriman House.
- Obstfeld, M., e Rogoff, K. (1996). *Foundations of international macroeconomics*. MIT Press.
- Oliveira, F. (1983). *A economia brasileira: Crítica à razão dualista*. Brasiliense.
- Ozturk, M., Toroslu, I. H., e Fidan, G. (2016). *Heuristic based trading system on Forex data using technical indicator rules*. *Applied Soft Computing*, 43, 170–186.
- Pardo, R. (2011). *The Evaluation and Optimization of Trading Strategies (2nd ed.)*. John Wiley e Sons.
- Park, C.-H., e Irwin, S. H. (2007). *What do we know about the profitability of technical*

analysis? *Journal of Economic Surveys*, 21(4), 786–826.

Petropoulos, F., Hyndman, R. J., Bergmeir, C., e others. (2022). *Model combinations through revised base rates*. *International Journal of Forecasting*, 38(4), 1261–1276.

Politis, D. N., e Romano, J. P. (1994). *The stationary bootstrap*. *Journal of the American Statistical Association*, 89(428), 1303–1313.

Ponsi, E. (2007). *Forex Patterns e Probabilities*. Wiley.

Pring, M. J. (2002). *Technical analysis explained: The successful investor's guide to spotting investment trends and turning points (4th ed.)*. McGraw-Hill.

Qian, B., e Rasheed, K. (2004). *Hurst exponent and financial market predictability*. In *Proceedings of the 2nd IASTED International Conference on Financial Engineering and Applications* (pp. 203–209). Cambridge, MA: IASTED.

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufmann.

Quinlan, J. R. (1996). *Improved use of continuous attributes in C4.5*. *Journal of Artificial Intelligence Research*, 4, 77–90.

Raschka, S. (2017). *Python machine learning (2^a ed.)*. Packt Publishing.

Rasekhschaffe, K. C., e Jones, R. C. (2019). *Machine learning for stock selection*. *Financial Analysts Journal*, 75(3), 70–88.

Reddy, B. P. (2019). *The role of feature engineering in machine learning: Techniques, challenges, and automation*. *IRE Journals*, 2(7), 7–13.

Rego, P. (2014). *Medidas de desempenho ajustadas ao risco: uma análise teórica e prática (Dissertação de Mestrado)*. Instituto Politécnico do Porto.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., e Müller, M. (2011). *pROC: an open-source package for R and S+ to analyze and compare ROC curves*. *BMC Bioinformatics*, 12, 77.

Rockafellar, R. T., e Uryasev, S. (2000). *Optimization of conditional value-at-risk*. *Journal of Risk*, 2(3), 21–41.

Rogers, L. C. G., & Satchell, S. E. (1991). *Estimating variance from high, low and closing prices*. *The Annals of Applied Probability*, 1(4), 504–512.

Roh, Y., Heo, G., e Whang, S. E. (2019). *A survey on data collection for machine learning: A big data-AI integration perspective*. *IEEE Transactions on Knowledge and Data*

Engineering, 33(4), 1328–1347.

Siddiqi, N. (2006). *Credit Risk Scorecards*. Wiley.

Siddiqi, N. (2012). *Credit risk scorecards: Developing and implementing intelligent credit scoring*. Wiley.

Ślepaczuk, R., e Zakrzewski, G. (2009). *High-Frequency and Model-Free Volatility Estimators*. Working Paper No. 13/2009, UniveIFRty of Warsaw.

Sortino, F. A. (1991). *Performance measurement in a downside risk framework*. *Journal of Investing*, 1(2), 59–64.

Sortino, F. A., e Van der Meer, R. (1991). *Downside risk*. *Journal of Portfolio Management*, 17(4), 27–31.

Stock, J. H., e Watson, M. W. (2015). *Introduction to econometrics (3rd ed.)*. Pearson.

Tan, P. N., Steinbach, M., Karpatne, A., e Kumar, V. (2021). *Introduction to Data Mining (2nd ed.)*. Pearson.

Thomas, L. C., Crook, J., e Edelman, D. (2017). *Credit Scoring and Its Applications*. SIAM.

Thu, T. N. T., e Xuan, V. D. (2018). Using support vector machine in Forex predicting. *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, 1–5.

Tibshirani, R. (1996). *Regression shrinkage and selection via the lasso*. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.

Tiwari, A. K., Nasreen, S., e Shahbaz, M. (2024). *Testing the efficient market hypothesis in foreign exchange markets: Evidence from nonlinear and multifractal approaches*. *Finance Research Letters*, 62, 104876.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. 3rd ed. Hoboken, NJ: Wiley.

Tsuji, C. (2015). *The dynamic interactions between the stock and foreign exchange markets: Evidence from Japan*. *International Journal of Financial Studies*, 3(3), 220–236.

Uzun, I., Lobachev, M., Kharchenko, V., Schöler, T., e Lobachev, I. (2024). *Candlestick pattern recognition in cryptocurrency price time-series data using rule-based data analysis methods*. *Computation*, 12(7), 132.

Van Buuren, S. (2018). *Flexible Imputation of Missing Data (2nd ed.)*. CRC Press.

Verdelhan, A. (2018). *The share of systematic risk in exchange rates*. *Journal of Finance*,

73(1), 375–418.

Wilder, J. W. (1978). *New Concepts in Technical Trading Systems*. Trend Research.

Williams, L. (1979). *How I made one million dollars last year trading commodities*. Windsor Books.

Yu, L., Wang, S., e Lai, K. K. (2005). *A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates*. *Computers e Operations Research*, 32(11), 2523–2541.

Zhang, Z., e Khushi, M. (2020). *GA-MSSR: Genetic Algorithm Maximizing Sharpe and Sterling Ratio Method for RoboTrading (preprint)*.

APÊNDICE

Apêndice A: Gráficos

Evolução Temporal

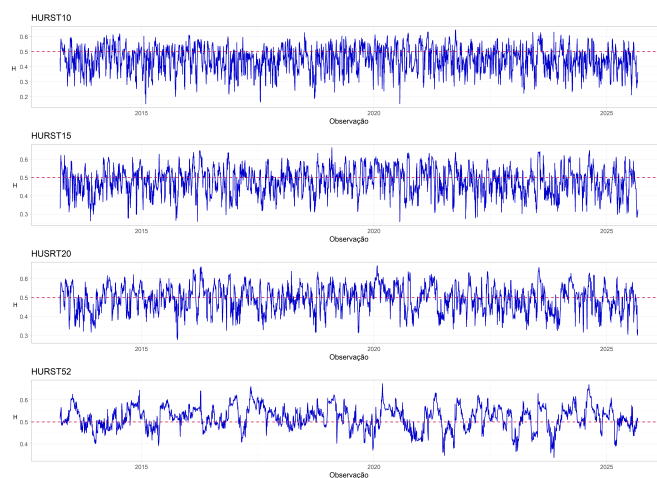


Figura 4.34: Evolução temporal do coeficiente de HURST por taxa de variação do USD/JPY.

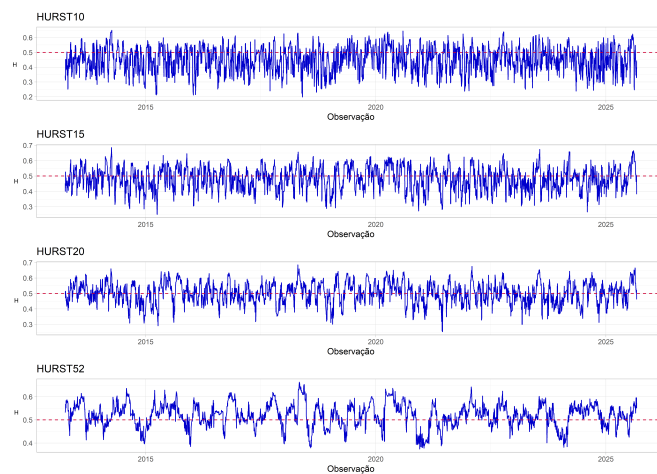


Figura 4.35: Evolução temporal do coeficiente de HURST por taxa de variação do GBP/USD.

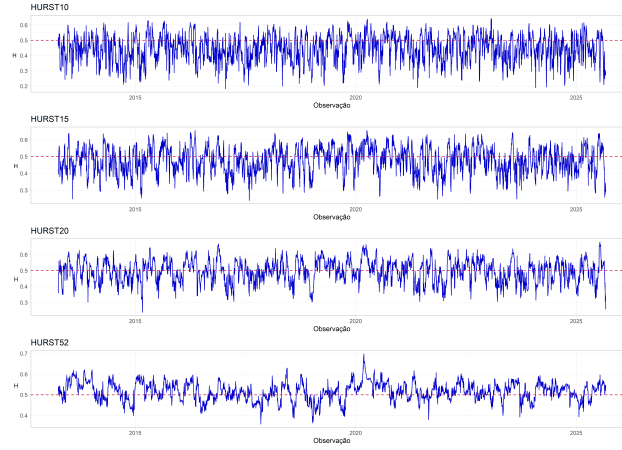


Figura 4.36: Evolução temporal do coeficiente de HURST por taxa de variação do EUR/USD.



Figura 4.37: Evolução temporal da probabilidade estacionariedade do USD/JPY por taxa de variação.

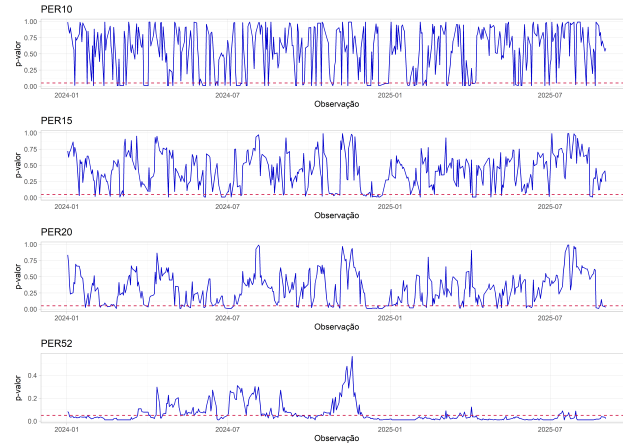


Figura 4.38: Evolução temporal da probabilidade estacionariedade do EUR/USD por taxa de variação.

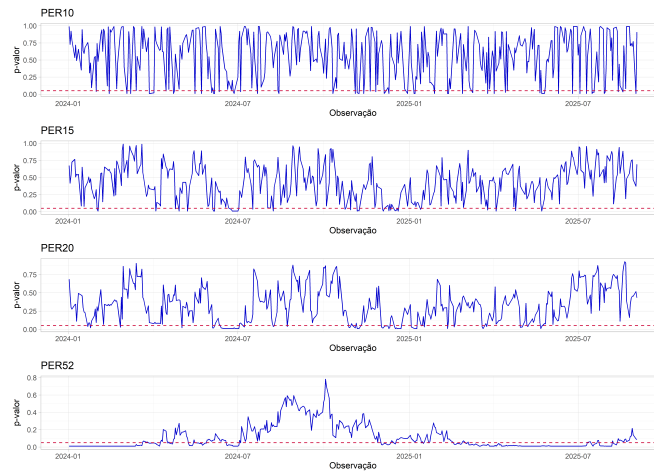
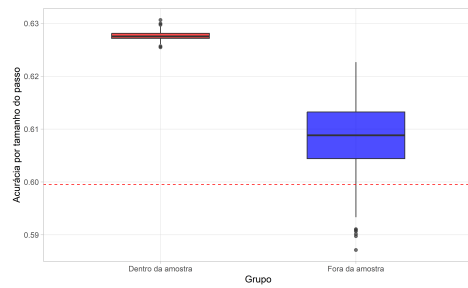
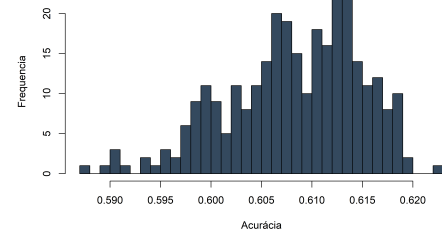


Figura 4.39: Evolução temporal da probabilidade estacionariedade do GBP/USD por taxa de variação.

Distribuição

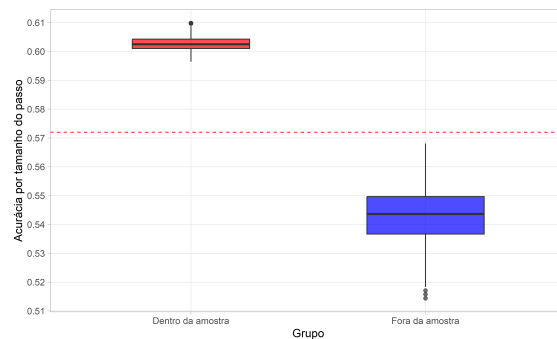


(a)

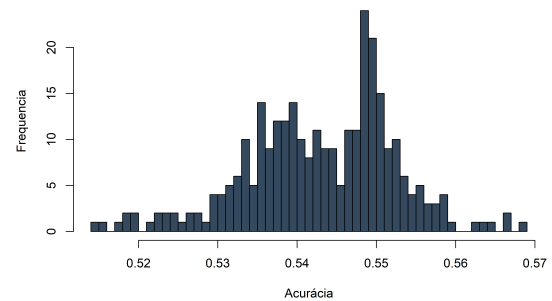


(b)

Figura 4.40: Desempenho da validação cruzada para a regressão logística ($L1$) no par EUR/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

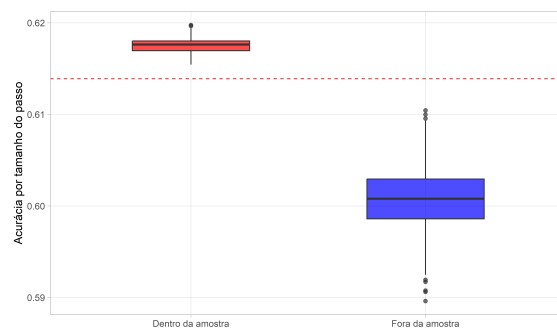


(a)

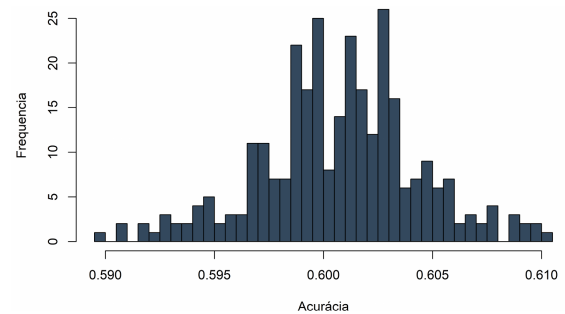


(b)

Figura 4.41: Desempenho da validação cruzada para a árvore de decisão no par EUR/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

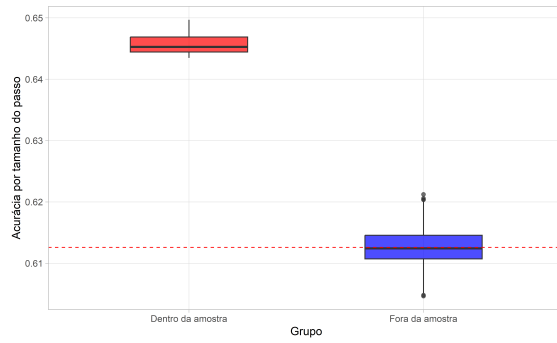


(a)

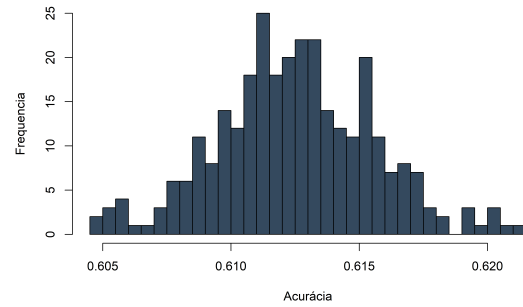


(b)

Figura 4.42: Desempenho da validação cruzada para a rede neural no par EUR/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

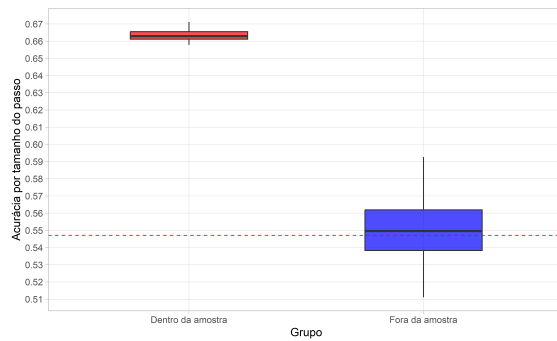


(a)

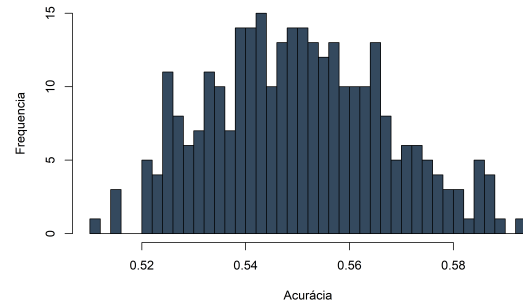


(b)

Figura 4.43: Desempenho da validação cruzada para a regressão logística ($L1$) no par GBP/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

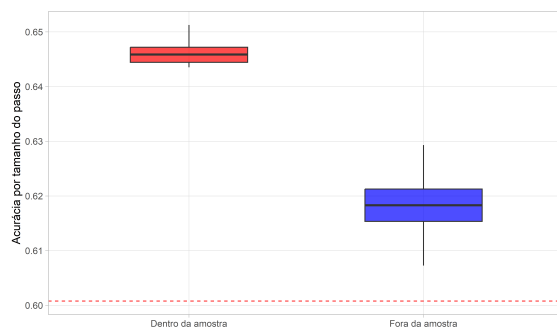


(a)

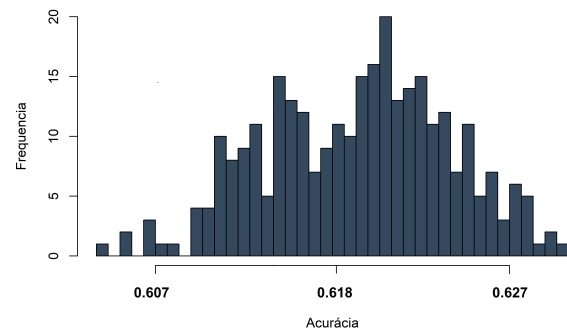


(b)

Figura 4.44: Desempenho da validação cruzada para a árvore de decisão ($L1$) no par GBP/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

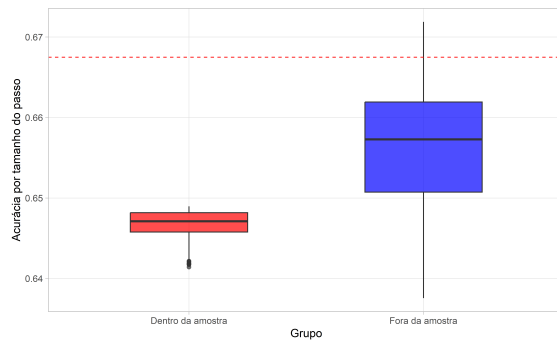


(a)

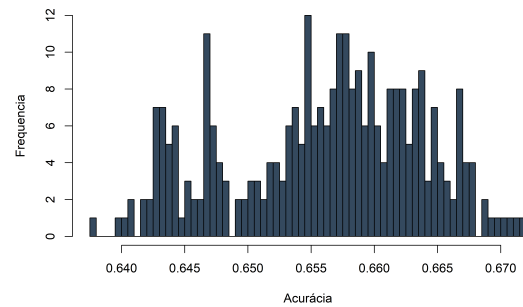


(b)

Figura 4.45: Desempenho da validação cruzada para a rede neural no par GBP/USD: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

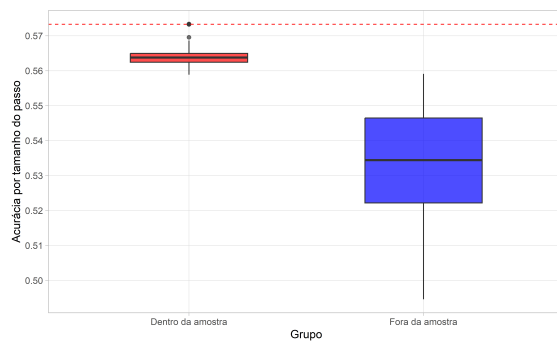


(a)

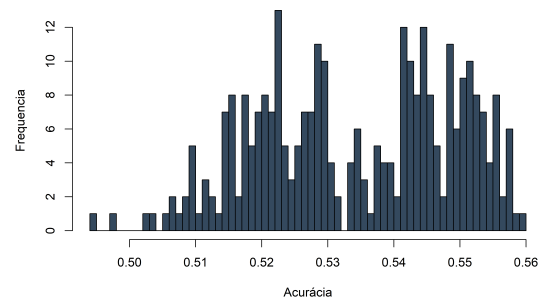


(b)

Figura 4.46: Desempenho da validação cruzada para a regressão logística ($L1$) no par USD/JPY: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

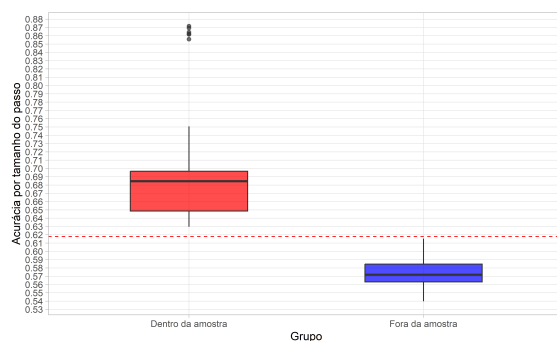


(a)

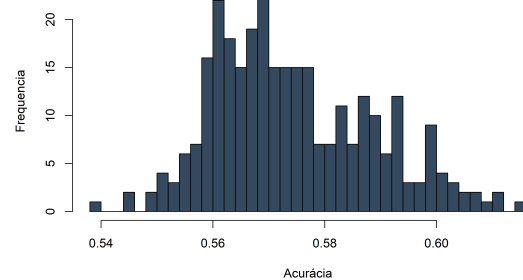


(b)

Figura 4.47: Desempenho da validação cruzada para a árvore de decisão ($L1$) no par USD/JPY: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.



(a)



(b)

Figura 4.48: Desempenho da validação cruzada para a rede neural no par USD/JPY: (a) Boxplot das acurácias médias por tamanho de passo dentro e fora da amostra; (b) Distribuição das acurácias médias fora da amostra.

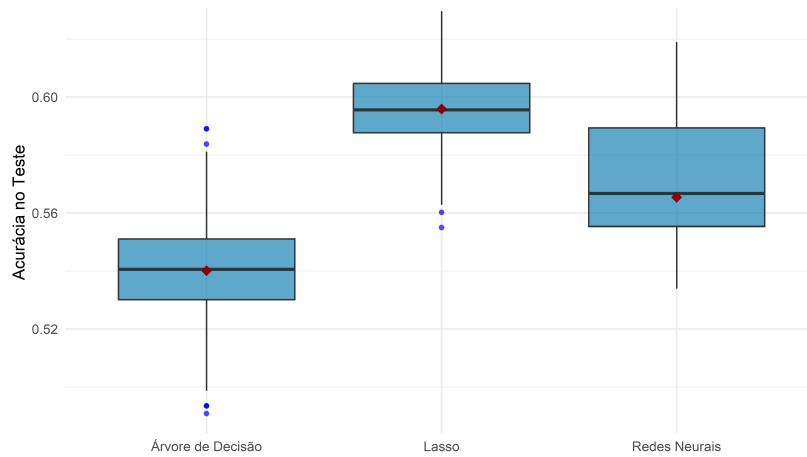


Figura 4.49: Distribuição das acurácias no bootstrap por blocos moveis dos algoritmos do par EUR/USD.

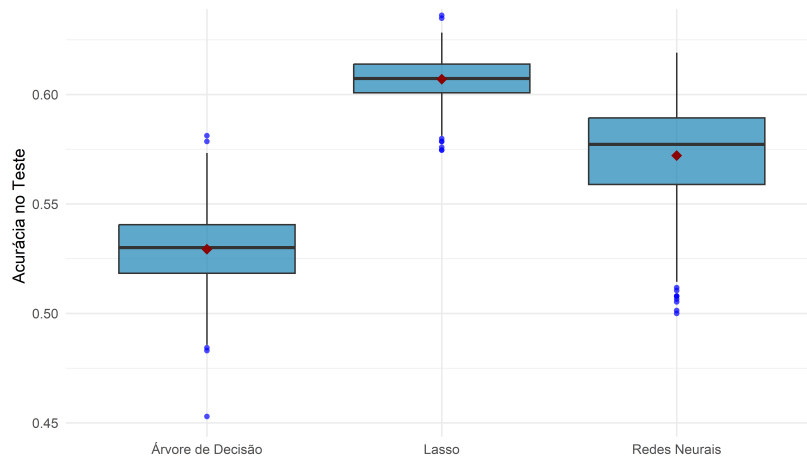


Figura 4.50: Distribuição das acurácias no bootstrap por blocos moveis dos algoritmos do par GBP/USD.

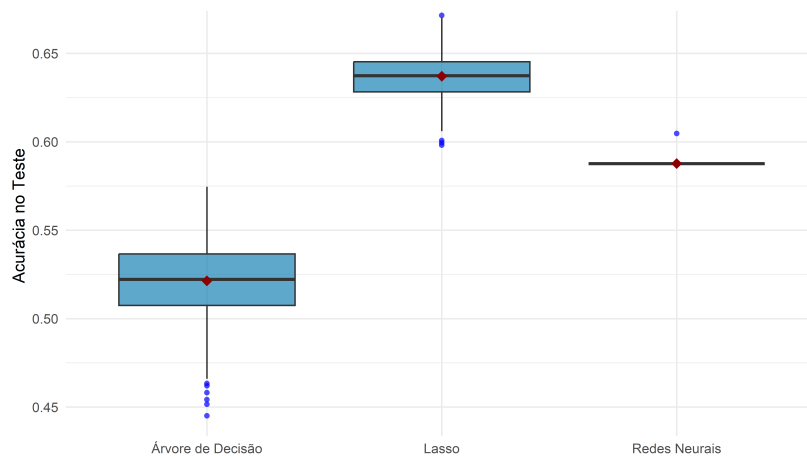


Figura 4.51: Distribuição das acurácias no bootstrap por blocos moveis dos algoritmos do par USD/JPY.

Heat Map

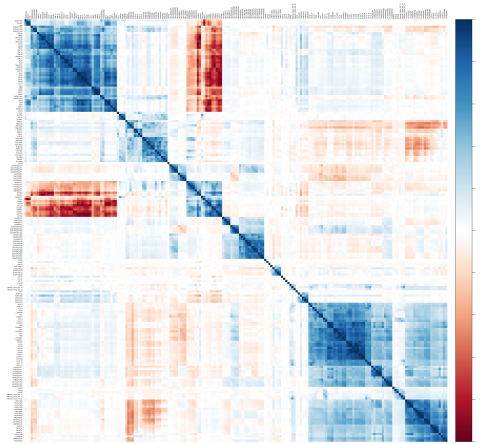


Figura 4.52: Heat Map da correlação entre as variáveis do par EUR/USD.



Figura 4.53: Heat Map da correlação entre as variáveis do par GBP/USD.

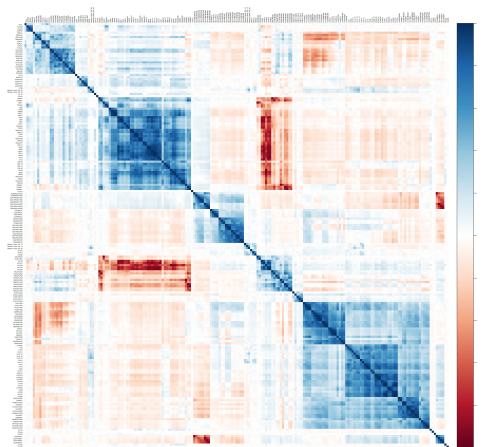


Figura 4.54: Heat Map da correlação entre as variáveis do par USD/JPY.

Árvore de decisão otimizada

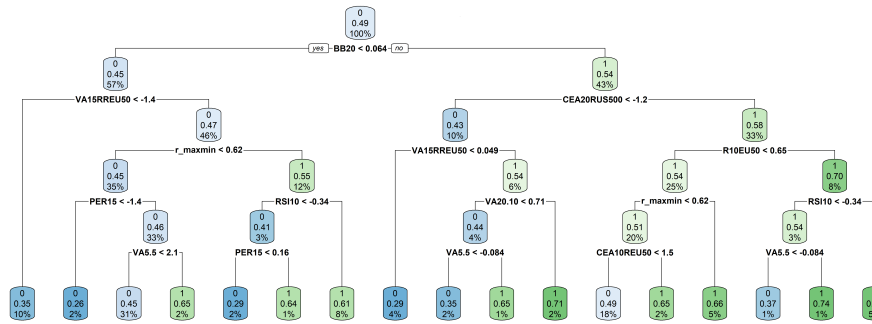


Figura 4.55: Árvore de decisão do par EUR/USD.

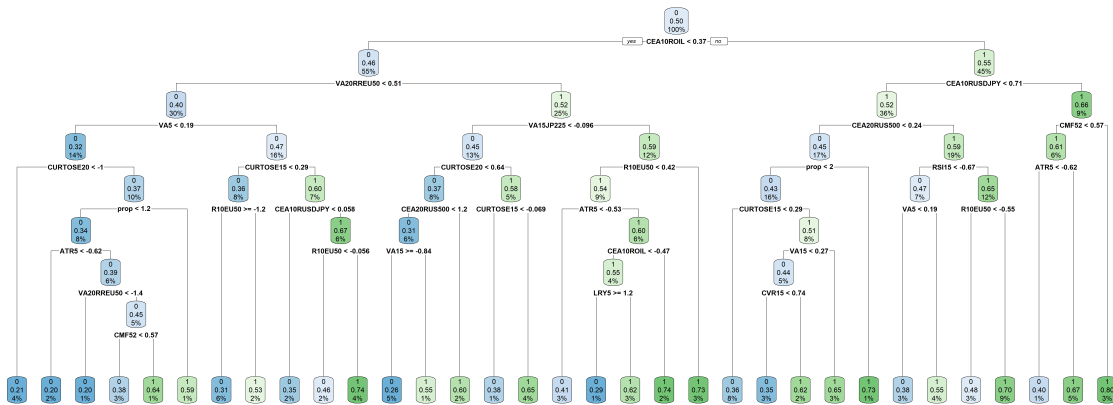


Figura 4.56: Árvore de decisão do par GBP/USD.

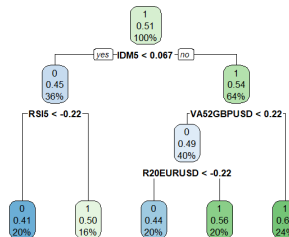


Figura 4.57: Árvore de decisão do par USD/JPY.

Retorno Acumulado Bruto



Figura 4.58: Curvas dos retornos acumulados dos EUR/USD.



Figura 4.59: Curvas dos retornos acumulados dos GBP/USD.

Apêndice B: Tabelas auxiliares

Tabela 4.15: Coeficientes da regressão logística penalizada via L1 otimizada para USDJPY, EURUSD e GBPUSD.

USDJPY		EURUSD		GBPUSD	
Variáveis	Coefficientes	Variáveis	Coefficientes	Variáveis	Coefficientes
r_maxmin	-0,112	VA150IL	0,045	(Intercept)	-0,007
OES	-0,081	WPR52	0,050	VA15	0,067
WPR52	-0,079	VA15JP225	0,055	Retorno10	0,070
RJP225	-0,078	VA5US500	0,058	LR5	0,072
CMF10	-0,075	RSI5	0,066	R10EU50	0,078
VA10,10	-0,065	Retorno	0,067	OMC52	0,101
EVRS5	-0,057	VA10GBPUSD	0,072	CVR15	0,119
VA15US500	-0,056	Retorno15	0,072	prop	0,123
VA15RREU50	-0,056	CEA10REU50	0,074	R5DXY	0,125
R520IL	-0,044	VA5,5	0,092	VA5	0,130
CEA20RUS500	-0,041	VA20,10	0,093	CURTOSE20	0,154
CEAZORDXY	-0,027	ASMTR5	0,097	CEA10RUSDJPY	0,162
R10EURUSD	-0,027	CEA20RUS500	0,099	CMF52	0,162
EVRS20	-0,027	RSI10	0,100	VA20RREU50	0,163
Retorno5	-0,025	BB20	0,107	ATR5	0,167
ROIL	-0,019	R10EU50	0,112	RSI15	0,171
R5110	-0,012	r_maxmin	0,116	CEA20RUS500	0,172
OE10	0,005	PER15	0,121	CEA10ROIL	0,177
R52EURUSD	0,008	VA15RREU50	0,138	R5EU50	0,189
Hurst RS_10	0,008	(Intercept)	.	VA15JP225	0,191
CEA15RJP225/20RJP225	0,016	CEA20ROIL	.	CURTOSE15	0,205
CMF15	0,021				
ASMTR15	0,022				
R5JP225	0,027				
(Intercept)	0,027				
VA52GBPUSD	0,038				
EVRS15	0,041				
Retorno	0,045				
REUSO	0,053				
R200IL	0,059				
CVR10	0,063				
CEA10REU50	0,063				
R20EU50	0,068				
EVRS52	0,075				
REURUSD	0,076				
CEA10RGBPUSD	0,079				
Hurst RS_15	0,091				
RGBPUSD	0,094				
BB5	0,094				
VA20JP225	0,102				
R10JP225	0,104				
VA20,10	0,104				
PE52	0,104				
VAS	0,106				
CVR20	0,115				
ASMTR20	0,116				
VASRREUSO	0,118				
R5EU50	0,119				
VAS2U5500	0,119				
IDM5	0,129				
R20EURUSD	0,134				
RSIS	0,170				
VA20EURUSD	0,206				

Tabela 4.16: Parâmetros Utilizados nas Estratégias de Negociação.

Par de Moedas	U_Prob	L_Prob	$H(u)$	$H(d)$	$I(k)$	VaR	$cost_op$
EUR/USD	0,5	0,5	0,5486	0,5086	39,346	-0,798	0,01
GBP/USD	0,5	0,5	0,5486	0,5086	39,346	-0,954	0,01
USD/JPY	0,5	0,5	0,5486	0,5086	39,346	-2,216	0,01