



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

João Felipe Morais Vasconcelos Sobral

Modelagem de Incerteza com Random Forest para Estratégias de Investimento em Ações

Recife

2025

João Felipe Morais Vasconcelos Sobral

Modelagem de Incerteza com Random Forest para Estratégias de Investimento em Ações

Trabalho apresentado ao Programa de Graduação em Ciências da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Ciências da Computação.

Área de Concentração: Inteligência Computacional

Orientador : Adriano Lorena Inácio de Oliveira

Recife

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Morais Vasconcelos Sobral, João Felipe.

Modelagem de Incerteza com Random Forest para Estratégias de
Investimento em Ações / João Felipe Moraes Vasconcelos Sobral. - Recife, 2025.
45 : il., tab.

Orientador(a): Adriano Lorena Inácio de Oliveira

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro de Informática, Ciências da Computação - Bacharelado,
2025.

Inclui referências, apêndices.

1. Aprendizagem de máquina. 2. Ciência de dados. 3. Mercado
financeiro. 4. Bolsa de valores. 5. Day Trade. 6. Random Forest. I. Lorena
Inácio de Oliveira, Adriano. (Orientação). II. Título.

000 CDD (22.ed.)

João Felipe Morais Vasconcelos Sobral

**Modelagem de Incerteza com Random Forest para Estratégias de Investimento
em Ações**

Trabalho de Conclusão de Curso
apresentado ao Curso de Graduação em
Ciências da Computação da Universidade
Federal de Pernambuco, como requisito
parcial para obtenção do título de
bacharel em Ciências da Computação.

Aprovado em: 10/04/2025

BANCA EXAMINADORA

Profa. Dr. Adriano Lorena Inácio de Oliveira (Orientador)

Universidade Federal de Pernambuco

Prof. Dr. Adiel Teixeira de Almeida Filho (Examinador Interno)

Universidade Federal de Pernambuco

AGRADECIMENTOS

À minha falecida mãe, Ana Beatriz, por, mesmo que brevemente, me incentivar a estudar, acompanhar e ensinar.

Ao meu pai, Marcelo Antônio, por também me incentivar e ser exigente com meu desempenho nos estudos.

À minha avó, Maria de Fátima, por sempre acreditar em mim e me apoiar ao máximo na busca pelos meus objetivos, não importando quantas vezes eles tenham mudado.

Ao meu avô, Joaquim, por me ensinar e ser exemplo dos frutos do trabalho, da dedicação e do esforço.

À minha falecida avó, Ana, por ser uma referência de amor, bondade e fé em Deus.

Aos meus irmãos, Renata, Marcelo, Marcela, Lucas, Maria Fernanda e Joaquim, por serem meus primeiros amigos e companheiros de vida.

À minha noiva, Beatriz, por sempre me apoiar, escutar e iluminar minha vida durante toda a trajetória da graduação.

Aos meus amigos do Centro de Informática — e além dele — Caio, Júlio Cesar e Matheus, por tantos conhecimentos trocados e projetos bem-sucedidos.

Aos meus amigos Ana Luiza, Bernardo, Carlos Alberto, Lucas, Nara e Pedro Emmanuel, por todo o apoio no cotidiano e nos momentos de descontração.

Ao meu professor e orientador, Adriano Lorena Inácio de Oliveira, por todo o apoio, ensinamentos e pela atenta revisão na construção deste trabalho.

A Deus, principalmente, por me dar o tom da vida, a capacidade de aprender e todo o restante.

É necessário que ao menos uma vez na vida você duvide, tanto quanto possível, de todas as coisas.

— René Descartes

RESUMO

A previsão precisa dos preços das ações no mercado financeiro pode gerar ganhos expressivos para investidores em curtos períodos de tempo. Este estudo investiga a aplicação de um limite de incerteza, calculado a partir do desvio padrão das previsões de um modelo Random Forest Regressor, como critério para decidir a realização de operações intradiárias na bolsa de valores. Experimentos foram conduzidos com seis ações do mercado brasileiro para avaliar o impacto dessa abordagem. Os resultados indicam que a inclusão da incerteza como fator decisório mitigou os riscos operacionais ao reduzir significativamente o Máximo Drawdown. Para quantificar esse impacto, foram utilizadas métricas como Taxa de Sucesso, Montante Final, Dias Operados e o próprio Máximo Drawdown, evidenciando a eficácia da metodologia na otimização das estratégias de investimento.

Palavras-chaves: Aprendizagem de máquina; Ciência de dados; Mercado financeiro; Bolsa de valores; Day Trade; Random Forest.

ABSTRACT

Accurate forecasting of stock prices in financial markets can yield significant gains for investors over short periods of time. This study investigates the use of an uncertainty threshold, calculated from the standard deviation of predictions made by a Random Forest Regressor, as a decision criterion for executing intraday trading operations on the stock exchange. Experiments were conducted using six stocks from the Brazilian market to evaluate the impact of this approach. The results indicate that incorporating uncertainty as a decision-making factor helped mitigate operational risks by significantly reducing the Maximum Drawdown. To quantify this impact, metrics such as Success Rate, Final Capital, Number of Trading Days, and Maximum Drawdown were employed, demonstrating the effectiveness of the methodology in optimizing investment strategies.

Keywords: Machine Learning; Data Science; Financial Market; Stock Exchange; Day Trade, Random Forest.

LISTA DE FIGURAS

Figura 1 – Diagrama do Sistema	16
Figura 2 – Representação dos períodos utilizados para treinamento, validação e teste, com suas respectivas durações em dias.	18
Figura 3 – Preço de Abertura da ação JBSS3 ao longo do tempo.	20
Figura 4 – Exemplo Desvio Padrão do Regressor Random Forest prevendo valor de alta de determinado dia.	22
Figura 5 – Distribuição da Incerteza em 40 amostras.	23
Figura 6 – Exemplo de intervalo de incerteza com limitação em 30%.	25
Figura 7 – Exemplo de Drawdowns.	29
Figura 8 – Preço de abertura das ações ao longo do tempo.	32
Figura 9 – Frequência das normalizações que obtiveram o melhor Final Score por ação.	33
Figura 10 – Resultado ABEV3 usando intervalo de incerteza.	36
Figura 11 – Melhor resultado ABEV3 sem utilizar incerteza.	36
Figura 12 – Resultado BBDC4 usando intervalo de incerteza.	37
Figura 13 – Melhor resultado BBDC4 sem utilizar incerteza.	37
Figura 14 – Resultado EGIE3 usando intervalo de incerteza.	38
Figura 15 – Melhor resultado EGIE3 sem utilizar incerteza.	38
Figura 16 – Resultado ITUB4 usando intervalo de incerteza.	39
Figura 17 – Melhor resultado ITUB4 sem utilizar incerteza.	39
Figura 18 – Resultado JBSS3 usando intervalo de incerteza.	40
Figura 19 – Melhor resultado JBSS3 sem utilizar incerteza.	40
Figura 20 – Resultado PETR4 usando intervalo de incerteza.	41
Figura 21 – Melhor resultado PETR4 sem utilizar incerteza.	41
Figura 22 – Falha no uso da incerteza.	44

LISTA DE CÓDIGOS

Código Fonte 1 – Seleção da melhor predição com base nos menores valores de MAPE	23
Código Fonte 2 – Criação de intervalos de incerteza	25

LISTA DE TABELAS

Tabela 1 – Colunas originais dos datasets.	18
Tabela 2 – Parâmetros para treinamento do modelo.	19
Tabela 3 – Colunas resultantes da Predição	24
Tabela 4 – Parâmetros da função de trading.	27
Tabela 5 – Melhores resultados com e sem normalização.	32
Tabela 6 – ABEV3	33
Tabela 7 – BBDC4	33
Tabela 8 – EGIE3	34
Tabela 9 – ITUB4	34
Tabela 10 – JBSS3	34
Tabela 11 – PETR4	34
Tabela 12 – Resultados ABEV3.	36
Tabela 13 – Resultados BBDC4.	37
Tabela 14 – Resultados EGIE3.	38
Tabela 15 – Resultados ITUB4.	39
Tabela 16 – Resultados JBSS3.	40
Tabela 17 – Resultados PETR4.	41

LISTA DE ABREVIATURAS E SIGLAS

MDD Máximo Drawdown

LISTA DE SÍMBOLOS

∈ Pertence

SUMÁRIO

1	INTRODUÇÃO	12
1.1	MERCADO FINANCEIRO	12
1.2	OBJETIVO DO TRABALHO	12
2	TRABALHOS RELACIONADOS	14
3	METODOLOGIA	16
3.1	VISÃO GERAL DO SISTEMA	16
3.2	DATASETS	17
3.3	SISTEMA	18
3.3.1	Preparo dos Dados	19
3.3.1.1	<i>Seleção de Colunas</i>	19
3.3.1.2	<i>Normalização</i>	20
3.3.1.3	<i>Cálculo da Incerteza</i>	21
3.3.2	Treinamento	23
3.4	SISTEMA DAY TRADE	24
3.4.1	Uso da Incerteza	25
3.4.2	Regras para entrada e saída de operações	26
3.4.3	Mecanismos de controle de risco	28
3.4.4	Máximo Drawdown	28
4	EXPERIMENTOS E RESULTADOS	30
4.1	CONFIGURAÇÕES DOS EXPERIMENTOS	30
4.2	FINAL SCORE	30
4.3	NORMALIZAÇÃO	31
4.4	INTERVALO DE INCERTEZA	33
4.5	RESULTADOS FINAIS	35
5	TRABALHOS FUTUROS	43
	REFERÊNCIAS	45

1 INTRODUÇÃO

1.1 MERCADO FINANCEIRO

No mundo dos investimentos, a compra de ações de empresas consolidadas é uma das formas mais comuns de alocação de capital. Além da possibilidade de rentabilidade por meio de dividendos, os investidores também buscam a valorização do ativo ao longo do tempo, obtendo lucro ao vendê-lo por um preço superior ao de compra.

Entre as diversas estratégias existentes no mercado, o Day Trade se destaca por sua abordagem de curto prazo, na qual operações de compra e venda são realizadas dentro de um único pregão. Esse método visa capturar pequenas variações de preço ao longo do dia para obter ganhos rápidos. No entanto, sua eficácia depende de uma análise precisa para identificar os momentos ideais de entrada e saída, tornando-se um desafio amplamente estudado na literatura acadêmica (DÍAZ; HANSEN; CABRERA, 2024).

Diante da necessidade de estratégias mais robustas para minimizar riscos e maximizar retornos, estudos recentes têm explorado o uso de modelos de aprendizado de máquina para aprimorar a tomada de decisão no mercado financeiro. Em especial, técnicas como Random Forest Regressor têm-se mostrado promissoras na previsão dos valores máximos e mínimos das ações, permitindo operações mais informadas e estratégias de mitigação de riscos (VIJH et al., 2020).

1.2 OBJETIVO DO TRABALHO

Este trabalho tem como objetivo principal aprimorar a aplicação do modelo Random Forest Regressor para previsões no mercado de ações, com foco no desenvolvimento e uso de uma métrica de incerteza como critério estratégico para a realização de operações intradiárias.

A proposta se baseia na continuidade de estudos anteriores que exploraram o uso de algoritmos de aprendizado de máquina para prever os valores máximos e mínimos de ações e avaliaram o impacto de estratégias de controle de risco na rentabilidade de operações de Day Trade. A partir desses fundamentos, esta pesquisa visa avançar na eficácia das decisões operacionais por meio da incorporação da incerteza estatística como elemento central da lógica de operação.

O diferencial está na inclusão de métricas de incerteza diretamente no processo de deci-

são. Além de fornecer uma previsão convencional sobre o movimento do ativo e seus valores extremos, o modelo também gera uma medida associada ao grau de confiança dessa previsão. Essa métrica permite que o sistema de trading avalie o nível de confiabilidade das estimativas antes de decidir operar.

Assim, em cenários onde a incerteza é elevada, o sistema pode optar por não executar operações, evitando negociações em contextos de baixa confiabilidade. Espera-se, com isso, reduzir a exposição a riscos desnecessários, preservar o capital e otimizar o desempenho das estratégias adotadas, contribuindo para uma gestão de risco mais eficiente e adaptável às variações do mercado.

2 TRABALHOS RELACIONADOS

A previsão de preços de ações é um tema recorrente na literatura acadêmica, com abordagens que vão desde métodos estatísticos tradicionais até técnicas avançadas de aprendizado de máquina. Este trabalho se baseia, principalmente, nas contribuições de Albuquerque Filho (FILHO, 2021) e Albuquerque Junior (JUNIOR, 2022), que exploraram o uso do modelo Random Forest Regressor para o desenvolvimento de estratégias de Day Trade.

Albuquerque Filho (FILHO, 2021) avaliou um sistema de Day Trading fundamentado inicialmente em Redes Neurais Artificiais (Artificial Neural Networks – ANN), inspirado no trabalho clássico de Martinez (MARTINEZ et al., 2009). No estudo de Martinez, foi proposto um modelo capaz de prever os valores máximos e mínimos diários de ações com base em indicadores técnicos e no valor de abertura do pregão. O treinamento da rede utilizou a técnica de janela deslizante (sliding window), que permitiu adaptação contínua às mudanças do mercado. As previsões foram integradas a um sistema automatizado de compra e venda com intervalos de 15 minutos, sendo validadas por métricas como lucro e drawdown em ativos como PETR4 e VALE4.

A partir dessa base, Albuquerque Filho (FILHO, 2021) propôs a substituição do modelo ANN por uma abordagem baseada em Random Forest Regressor, alcançando melhorias significativas na previsão dos valores extremos de preço. Uma das principais inovações foi a implementação de um mecanismo de correção dinâmica das previsões, com base no erro e desvio padrão obtidos durante a fase de validação. Essa adaptação resultou em um aumento no retorno anual, sem aumento no drawdown, evidenciando maior eficiência para aplicações em estratégias de investimento de curto prazo.

Dando continuidade a essa linha de pesquisa, Albuquerque Junior (JUNIOR, 2022) desenvolveu um sistema mais robusto de operações intradiárias com foco na gestão de risco. Utilizando o modelo Random Forest Regressor, seu trabalho investigou o impacto de diferentes combinações de indicadores técnicos na previsão de preços e testou ajustes nos mecanismos de controle operacional, como regras de compra, venda e stop-loss. A avaliação foi realizada em seis ações do mercado brasileiro ao longo de um ano, utilizando métricas como MAPE (Mean Absolute Percentage Error), retorno anual e máximo drawdown. O estudo também comparou os resultados com modelos consagrados, reforçando a viabilidade da abordagem e destacando seus principais desafios.

Complementando essa discussão, pesquisas mais recentes vêm abordando a importância da incerteza nas previsões de modelos de machine learning. Dutschmann e Baumann (DUTSCHMANN; BAUMANN, 2021), por exemplo, propuseram o uso da variabilidade das previsões entre as árvores de decisão de uma Random Forest como métrica de incerteza, utilizando o desvio padrão para estimar o grau de confiança nas previsões. Essa abordagem segue a mesma linha de Coulston (COULSTON et al., 2016), que também exploram a quantificação da incerteza com o objetivo de tornar os modelos mais robustos. Ambos os estudos reforçam o valor de incorporar a incerteza como uma variável crítica na tomada de decisão, especialmente em ambientes voláteis como o mercado financeiro.

Esses trabalhos fundamentam a proposta deste estudo, que busca aliar a capacidade preditiva do Random Forest Regressor à análise da incerteza para desenvolver uma estratégia mais seletiva e eficiente na realização de operações intradiárias, promovendo uma gestão de risco mais refinada e adaptável ao comportamento do mercado.

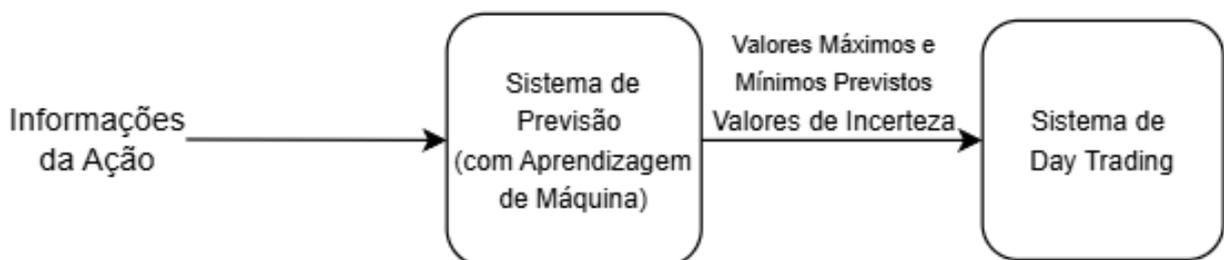
3 METODOLOGIA

Nesta sessão serão descritos com detalhes a ideação, motivação, implementação e avaliação dos métodos utilizados na pesquisa. O foco desse trabalho é avaliar de forma crítica o trabalho de Albuquerque Filho (FILHO, 2021) aplicando seu modelo de uma nova maneira para as operações Day Trading.

3.1 VISÃO GERAL DO SISTEMA

Esta seção apresenta o funcionamento do sistema proposto, detalhando o fluxo de informações, a aplicação do modelo de aprendizado de máquina e o mecanismo de tomada de decisão para a realização de operações intradiárias. O sistema foi desenvolvido com base na estrutura proposta por Albuquerque Filho (FILHO, 2021), que integra a modelagem preditiva com um sistema operacional de compra e venda de ações.

Figura 1 – Diagrama do Sistema



Fonte: Produzido pelo próprio Autor.

Como ilustrado na Figura 1, o sistema recebe dados históricos e atuais de uma ação selecionada, processando essas informações por meio de um modelo de aprendizado de máquina baseado no algoritmo Random Forest. O objetivo principal desse modelo é prever os valores máximos e mínimos do ativo ao longo do dia, permitindo identificar potenciais oportunidades de negociação com base na direção esperada do mercado.

Os valores previstos são, então, encaminhados para o módulo de operações intradiárias, responsável por executar ordens de compra e venda com base nos sinais gerados. O diferencial da abordagem proposta está na **incorporação de métricas de incerteza no processo decisório**. Além de fornecer previsões pontuais, o modelo também estima uma medida de incerteza associada a cada previsão, calculada a partir da variabilidade entre as árvores da floresta (desvio padrão das previsões individuais).

Essa métrica desempenha um papel central na estratégia: em cenários onde a incerteza é elevada — ou seja, quando o modelo apresenta baixa confiança na previsão — o sistema pode optar por não operar, evitando transações em contextos com alta volatilidade ou baixa previsibilidade. O limiar de incerteza é configurável, permitindo ao usuário ajustar o nível de tolerância ao risco de acordo com seu perfil ou estratégia.

Essa abordagem possibilita uma gestão de risco mais eficiente e adaptativa, reduzindo o número de operações potencialmente desfavoráveis e, conseqüentemente, o **Máximo Draw-down**. Com isso, o sistema busca maximizar os retornos das operações realizadas, atuando de forma seletiva e robusta frente às incertezas do mercado financeiro.

3.2 DATASETS

No estudo realizado por Albuquerque Junior (JUNIOR, 2022) foram utilizadas 6 ações, as quais também foram adotadas nesse trabalho: ABEV3, BBDC4, EGIE3, ITUB4, JBSS3 e PETR4. Essas ações são consolidadas e apresentam alto volume de transações diárias, o que tende a aumentar a liquidez e estabilizar o preço, reduzindo o risco de perdas decorrentes de atrasos na execução das ordens de compra e venda. Foram coletados dois conjuntos (Tabela 4) de dados com granularidades distintas: o primeiro, com registros diários, destinado ao treinamento do regressor, e o segundo, com registros a cada 15 minutos, para ser utilizado no Sistema de Day Trading.

Com o intuito de preservar a reprodutibilidade dos resultados e manter compatibilidade metodológica com o sistema de referência, optou-se por utilizar exatamente o mesmo conjunto de dados empregado por Albuquerque Junior (JUNIOR, 2022). Assim, foram adotados os seguintes períodos: de 2 de abril de 2018 a 3 de fevereiro de 2021 como conjunto de treinamento, de 4 de fevereiro de 2021 a 5 de abril de 2021 como conjunto de validação e, por fim, de 6 de abril de 2021 a 1 de junho de 2021 como conjunto de testes.

Diferentemente do uso convencional, o conjunto de validação aqui não tem o propósito de medir a eficácia do sistema, mas sim de criar um intervalo de incerteza, que será utilizado pelo Sistema de Day Trading no conjunto de teste. Construir esse intervalo a partir do conjunto de treinamento não seria adequado, pois os dados já foram utilizados para o treinamento do modelo. Da mesma forma, calculá-lo diretamente sobre o conjunto de teste configuraria um caso de vazamento de dados, ainda que de forma indireta. O cálculo da incerteza será detalhado posteriormente em uma seção específica dedicada ao tema.

Figura 2 – Representação dos períodos utilizados para treinamento, validação e teste, com suas respectivas durações em dias.



Fonte: Produzido pelo próprio Autor.

Tabela 1 – Colunas originais dos datasets.

Coluna	Significado
DATE	Data a que os dados se referem.
TIME	Momento que os dados foram coletados, esta coluna só está presente no dataset de operações.
OPEN	Preço de abertura da ação, corresponde à primeira operação realizada no dia, a partir das 10:00h. Pode ser próxima ao valor de fechamento do último dia de operação.
HIGH	Maior preço do ativo no dia.
LOW	Menor preço do ativo no dia.
CLOSE	Preço de fechamento do dia, valor corresponde à última operação. Operações terminam às 17:00h.
TICKVOL	Volume de operações, não leva em conta o volume de ações de cada, apenas a quantidade total de operações.
VOL	Volume de ações negociadas no dia.
SPREAD	Valor fracionário oriundo do desdobramento das ações. Ocorre quando uma empresa divide suas ações com finalidade de aumentar a circulação de papeis e diminuir o preço unitário.

Fonte: Produzido pelo próprio autor.

3.3 SISTEMA

Esse sistema tem como objetivo prever os valores máximos e mínimos que um determinado ativo irá alcançar e também calcular a incerteza dessas previsões de forma individual. Os resultados dele são usados em seguida pelo Sistema de Day Trading para realizar as operações.

3.3.1 Preparo dos Dados

3.3.1.1 Seleção de Colunas

Para o treinamento do modelo, foi utilizado um dataset com registros diários, excluindo as colunas VOL, TICKVOL e SPREAD por não fornecerem informações relevantes. Paralelamente, foram adicionadas colunas derivadas de indicadores técnicos, baseadas em dados anteriores, seguindo a premissa de que esses indicadores possuem potencial preditivo para valores futuros, especialmente em operações de curto prazo, conforme sugerido por Vanstone e Finnie (2009).

Baseado nisto, escolhemos dois indicadores clássicos, o EMA (Exponential Moving Average) e Bandas de Bollinger (BB). Na tabela 2 estão descritos todos os parâmetros com descrições, totalizando 33 colunas.

Tabela 2 – Parâmetros para treinamento do modelo.

Quantidade de Parâmetros	Nome da Coluna $X \in 1,2,3,4,5$	Descrição
1	open	Valor de abertura do dia corrente
2	EMA_high EMA_low	MME do Menor e Maior preço dos 5 últimos dias.
2	EMA_open EMA_close	MME do preço de Abertura e Fechamento dos 5 últimos dias.
4	UB_high LB_high UB_low LB_low	BB do Maior e Menor preço dos 5 últimos dias.
2	UB_open LB_open	BB dos preços de Abertura dos 5 últimos dias.
2	UB_close LB_close	BB dos preços de Fechamento dos 5 últimos dias.
10	d-X_high d-X_low	Menor e Maior preço dos 5 últimos dias.
10	d-X_open d-X_close	Preço de Abertura e Fechamento dos 5 últimos dias.

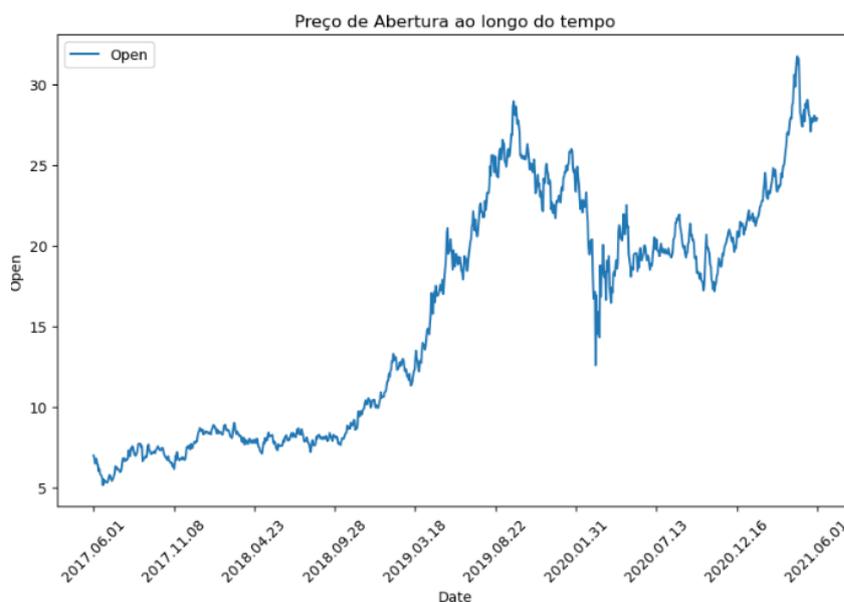
Fonte: Produzida pelo próprio autor baseada em Albuquerque Filho (FILHO, 2021).

3.3.1.2 Normalização

Considerando a variabilidade nos preços das ações ao longo dos anos, é essencial aplicar técnicas de normalização para garantir que o modelo de Random Forest Regressor não seja afetado por diferenças de escala nos dados. Algumas dessas variáveis passaram por variações significativas ao longo do tempo, o que pode levar a uma tendência de prever valores mais baixos se não forem feitos os ajustes necessários.

Para abordar esse desafio, escolhemos três normalizadores: Standard Scaler, MinMax Scaler e Power Scaler. Cada um desses normalizadores tem seus pontos fortes:

Figura 3 – Preço de Abertura da ação JBSS3 ao longo do tempo.



Fonte: Produzido pelo próprio Autor.

Standard Scaler

- **Padronização:** O Standard Scaler é útil para padronizar os dados, tornando-os mais fáceis de trabalhar com o Random Forest Regressor.
- **Melhora a convergência:** A padronização pode ajudar a melhorar a convergência do modelo, tornando-o mais eficiente.
- **Trabalha bem com dados contínuos:** O Standard Scaler é particularmente útil com dados contínuos, que são comuns em problemas de regressão.

MinMax Scaler

- **Limita os dados:** O MinMax Scaler é útil para limitar os dados em uma faixa específica, o que pode ajudar a prevenir problemas de escala com o Random Forest Regressor.
- **Preserva a relação entre os dados:** O MinMax Scaler preserva a relação entre os dados, o que é importante para o desempenho do modelo.
- **Trabalha bem com dados categóricos:** O MinMax Scaler pode ser útil com dados categóricos, desde que sejam codificados corretamente.

Power Scaler (ou Yeo-Johnson Transformer)

- **Trabalha bem com dados não gaussianos:** O Power Scaler é útil para trabalhar com dados que não seguem uma distribuição gaussiana, o que é comum em problemas de regressão.
- **Melhora a normalidade dos dados:** A transformação aplicada pelo Power Scaler pode ajudar a melhorar a normalidade dos dados, o que é importante para o desempenho do modelo.
- **Trabalha bem com dados contínuos:** O Power Scaler é particularmente útil com dados contínuos, que são comuns em problemas de regressão.

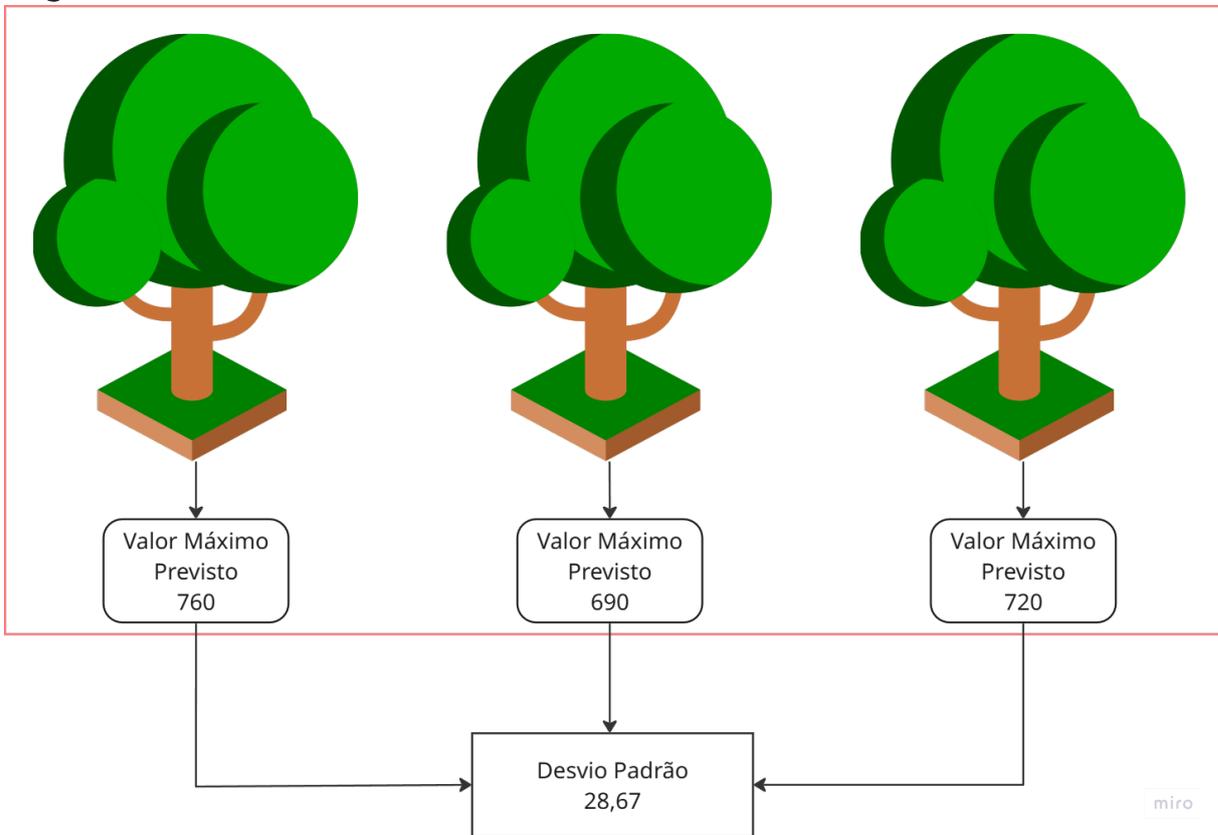
3.3.1.3 Cálculo da Incerteza

O modelo Random Forest Regressor opera mediante a combinação de múltiplas Árvores de Decisão, cada uma gerando uma previsão individual. Ao executar o modelo para prever o conjunto de testes, temos X árvores realizando uma regressão para estimar os valores de alta e baixa do dia. Com base nessa abordagem, foi calculado o desvio padrão das previsões geradas por todas as árvores da floresta para os valores mínimo e máximo estimados de cada dia (Figura 4). A média entre esses dois desvios padrão, conforme expressa na equação 3.1, é denominada incerteza diária. Essa métrica é independente dos valores reais observados, sendo determinada exclusivamente com base nas variações internas do modelo preditivo.

Esta metodologia segue o princípio proposto por estudos como o de Coulston (COULSTON et al., 2016) e Dutschmann e Baumann (DUTSCHMANN; BAUMANN, 2021), que utilizam a variabilidade das previsões das árvores individuais para quantificar a incerteza associada às predições agregadas do modelo. Ao calcular o desvio padrão das previsões das árvores, estamos

Figura 4 – Exemplo Desvio Padrão do Regressor Random Forest prevendo valor de alta de determinado dia.

Regressor Random Forest



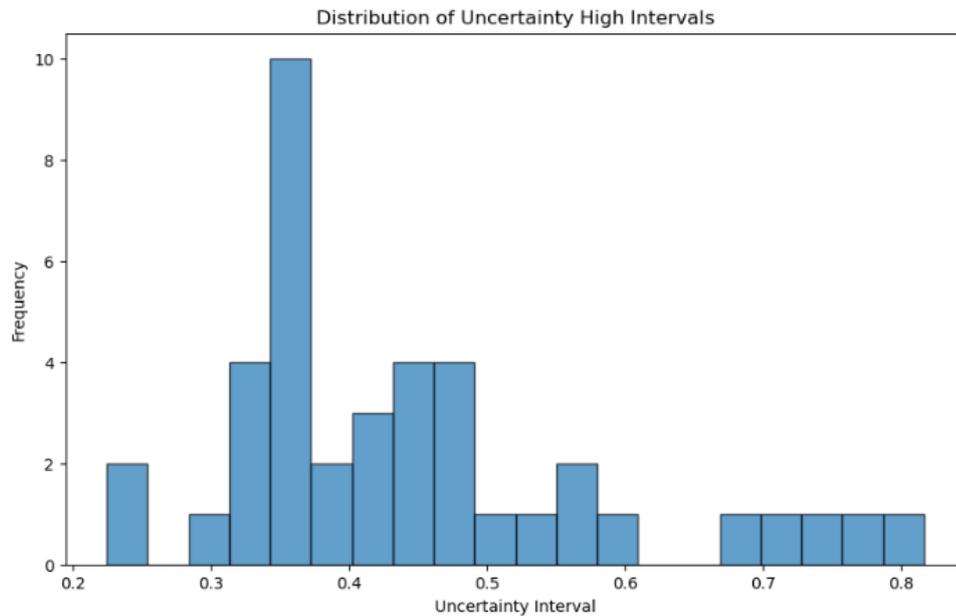
Fonte: Produzido pelo próprio Autor.

essencialmente estimando a dispersão da previsão do modelo e, portanto, a confiança que o modelo tem em suas previsões para cada dia.

Além da incerteza diária calculada anteriormente, é necessário estabelecer um valor de referência para qualificar a incerteza como alta ou baixa. Para isso, foi utilizada a abordagem de validação, que consiste em utilizar um conjunto de validação como parâmetro de referência para avaliar as incertezas futuras do conjunto de testes. Essa abordagem permite estabelecer um limite de tolerância para a incerteza, possibilitando a qualificação da incerteza como alta ou baixa em relação ao valor de referência.

$$\text{incerteza}_{\text{diária}} = \frac{\text{desvio padrão}_{\text{máximo}} + \text{desvio padrão}_{\text{mínimo}}}{2} \quad (3.1)$$

Figura 5 – Distribuição da Incerteza em 40 amostras.



Fonte: Produzido pelo próprio Autor.

3.3.2 Treinamento

O treinamento do modelo foi realizado utilizando a divisão do conjunto de dados diário, conforme descrito na Seção 3.2. Para cada normalizador, incluindo a opção sem normalização, foram realizadas quatro execuções, todas utilizando sementes aleatórias e 300 estimadores. O modelo mais adequado foi selecionado com base no Erro Médio Absoluto Percentual (MAPE), calculado a partir das previsões e valores reais tanto para as máximas quanto para as mínimas diárias. Nessas execuções, mantiveram-se fixos os hiperparâmetros do modelo, e os resultados de previsão — juntamente com os valores de incerteza — eram atualizados somente quando apresentavam redução do MAPE em ambos os alvos de previsão. Esse procedimento permitiu selecionar, dentre as repetições com os mesmos parâmetros, aquela com melhor desempenho.

Código Fonte 1 – Seleção da melhor previsão com base nos menores valores de MAPE

```

1 if errorMax[1] < best_errorMax[1] and errorMin[1] < best_errorMin[1]:
    best_y_pred = y_pred
3     best_uncertainty_valid = uncertainty_valid
    best_uncertainty_test = uncertainty_test
5     best_errorMax = errorMax
    best_errorMin = errorMin

```

Fonte: Produzido pelo próprio Autor.

Tabela 3 – Colunas resultantes da Predição

Coluna	Significado
date	Data a que os dados se referem.
high_pred	Previsão do maior preço do dia.
low_pred	Previsão do menor preço do dia.
uncertainty_mean	Média da incerteza da previsão dos valores de alta e baixa.
uncertainty_valid_mean	Média da incerteza do conjunto de validação para ser usado como referência.
high	Maior preço real do dia.
low	Menor preço real do dia.

Fonte: Produzido pelo próprio autor.

3.4 SISTEMA DAY TRADE

Este sistema tem como objetivo simular operações de compra e venda de ações em uma corretora, utilizando um conjunto de dados com registros a cada 15 minutos. A cada intervalo, o sistema avalia as condições do mercado e decide entre comprar, vender ou manter a posição inalterada. A metodologia adotada segue a abordagem proposta por Albuquerque Filho (FILHO, 2021), que, por sua vez, se baseou no trabalho de Martinez (MARTINEZ et al., 2009).

Como proposto por Chande (CHANDE, 2001), que sugere que um sistema de negociação deve possuir 3 partes principais:

1. Regras de compras e vendas;
2. Mecanismos de controle de riscos;
3. Gerenciamento de capital.

Para a simulação das operações, assumimos a ausência de taxas para compra e venda, condição oferecida por diversas corretoras. Além disso, consideramos um impacto negligenciável do slippage, uma vez que as transações envolvem lotes inteiros de ações. O capital inicial foi fixado em R\$500.000,00, permitindo que o sistema realize um maior número de operações e utilize um volume mais significativo de capital ao longo dos testes.

3.4.1 Uso da Incerteza

Conforme descrito na Seção 3.3.1.3, a incerteza é definida como a dispersão dos resultados das árvores de regressão utilizadas na predição dos valores mínimos e máximos de uma ação em um determinado dia. Para obter um referencial de incerteza média, foram consideradas as previsões dos 40 dias pertencentes ao conjunto de validação.

Com base nesses valores médios, construiu-se um intervalo com limites definidos pela maior e menor incerteza observada. Esse intervalo foi dividido em 20 segmentos de igual tamanho, escolha motivada pela facilidade de uso em percentuais, pela maior liberdade na definição do nível de segurança das operações e pela maior granularidade na análise dos resultados. A Figura 6 ilustra essa segmentação.

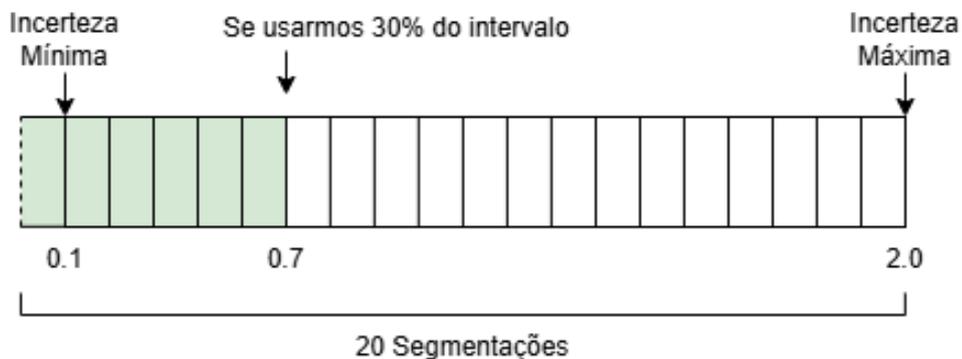
Código Fonte 2 – Criação de intervalos de incerteza

```
# Obtendo os valores minimo e maximo da incerteza para o dataframe atual
2 min_uncertainty = df_pred["uncertainty_valid_mean"].min()
  max_uncertainty = df_pred["uncertainty_valid_mean"].max()
4
# Definindo um passo dinamico (5% da diferenca total)
6 steps = np.linspace(min_uncertainty, max_uncertainty, num=20)
  steps = np.append(steps, 0)
```

Fonte: Produzido pelo próprio Autor.

Na execução do sistema, os limites desses segmentos permitiram a utilização de percentuais variáveis — de 5% a 100% do intervalo —, proporcionando um controle de risco mais flexível e alinhado ao nível de segurança desejado.

Figura 6 – Exemplo de intervalo de incerteza com limitação em 30%.



Fonte: Produzido pelo próprio Autor.

Durante a execução do Sistema de Day Trade, ao longo dos dias do conjunto de testes, a incerteza diária é avaliada conforme representado na Figura 3.1. Caso esse valor exceda o limite

definido pelo usuário (Figura 6), o sistema não realizará operações naquele dia, classificando-o como um dia de risco.

3.4.2 Regras para entrada e saída de operações

Seguindo o trabalho de Albuquerque Filho (FILHO, 2021), as regras que determinam os momentos de entrada e saída em operações de day trade são fundamentais para o funcionamento de um sistema de negociações, pois definem como as transações serão executadas. Na simulação realizada, a cada intervalo de 15 minutos, o sistema analisa o valor de fechamento do ativo e verifica se ele ultrapassa o limite superior ou inferior previsto, acionando uma operação quando essa condição é atendida. Dessa forma, é possível que, em determinados dias, nenhuma operação seja realizada caso o preço de fechamento permaneça dentro da faixa predita ao longo de todas as janelas de 15 minutos.

A escolha desse intervalo de tempo foi condicionada pelas características do dataset utilizado. No entanto, o impacto da variação desse período sobre os resultados do sistema é uma questão que pode ser explorada em estudos futuros.

É importante perceber que há uma regra intrínseca a este tipo de operação que será modelada no sistema: Toda operação de entrada (seja ela de compra ou de venda) deve ser seguida de uma de saída até o final do dia corrente, mesmo que resulte em prejuízo. Ou seja, se uma operação de entrada foi iniciada e não atingiu as condições definidas para uma saída, a saída será realizada nos últimos momentos do dia.

Seja \min_{pred} o valor mínimo predito pelo modelo de predição, \max_{pred} o valor máximo, fecho o valor de fechamento da janela de 15 minutos, todas estas restrições mencionadas anteriormente podem ser modeladas pelas seguintes regras:

1. Comprar quando $\text{fecho} \leq \min_{\text{pred}}$, vender quando $\text{fecho} \geq \max_{\text{pred}}$;
2. Comprar quando $\text{fecho} \leq \min_{\text{pred}}$, vender no último minuto do dia;
3. Comprar quando $\text{fecho} \leq \min_{\text{pred}}$, vender no preço de stop-loss;
4. Vender quando $\text{fecho} \geq \max_{\text{pred}}$, comprar quando $\text{fecho} \leq \min_{\text{pred}}$;
5. Vender quando $\text{fecho} \geq \max_{\text{pred}}$, comprar no último minuto do dia;
6. Vender quando $\text{fecho} \geq \max_{\text{pred}}$, comprar no preço de stop-loss;

Para aumentar a flexibilidade do sistema e permitir operações mais dinâmicas, foi implementado um critério baseado no desvio padrão aplicado aos valores mínimo e máximo previstos. Dessa forma, a execução das operações não exige que o preço atinja exatamente os limites preditivos, mas sim uma faixa ajustável em torno desses valores.

Tabela 4 – Parâmetros da função de trading.

Parâmetro	Significado
df	Dataset com resultado das previsões e incertezas.
dailyDf	Dataset que representa a variação dos preços das ações a cada 15 minutos.
money	Montante inicial disponível para as operações.
stopLoss	Define o limite percentual de perda aceitável em uma operação, encerrando-a automaticamente quando ultrapassado.
std	Ajusta as previsões do modelo ao aplicar uma margem percentual, tornando os limites de compra e venda mais sensíveis e influenciando diretamente a quantidade de operações executadas
useUncertainty	Valor booleano que indica se o sistema deve considerar o limiar de incerteza.
uncertaintyThreshold	Define um limite superior para a incerteza, fazendo com que o dia seja ignorado caso esse valor seja ultrapassado.

Fonte: Produzido pelo próprio autor.

Esse desvio padrão atua como uma margem de tolerância, reduzindo a rigidez do modelo e permitindo que as operações sejam iniciadas de forma mais responsiva às variações do mercado. A magnitude do desvio padrão é definida pelo usuário como um parâmetro configurável, permitindo adaptar a estratégia de negociação conforme diferentes perfis de risco e cenários de volatilidade.

A inclusão desse critério visa minimizar a chance de oportunidades perdidas devido a pequenas flutuações que poderiam impedir uma operação caso o preço não atingisse exatamente os valores projetados. Além disso, essa abordagem pode contribuir para uma maior estabilidade do sistema, reduzindo o impacto de oscilações momentâneas e tornando as decisões de entrada e saída mais robustas.

3.4.3 Mecanismos de controle de risco

Além da verificação da incerteza para não operação no dia como descrito na seção 3.4.1, usamos outros mecanismos para evitar perdas e delimitar limites. Sistemas de day trading normalmente possuem taxas de acertos similares às de erro, porém os volumes nos momentos de acertos tendem a ser maiores que nos de erro.

Visando evitar perdas, utilizamos o stop-loss, um mecanismo amplamente empregado em estratégias de negociação automatizada para limitar o risco de operações. O stop-loss é um valor predefinido que, quando atingido pelo preço do ativo, aciona automaticamente a saída da posição, evitando que perdas significativas ocorram devido a movimentos inesperados do mercado.

No contexto deste trabalho, o stop-loss foi implementado de forma a encerrar uma operação sempre que o preço da ação se movimentasse contra a direção esperada além de um limite previamente estabelecido. O valor utilizado para o stop-loss foi de 2%, o que significa que operações de compra eram encerradas automaticamente em caso de queda superior a 2%, enquanto operações de venda eram encerradas em caso de alta superior a esse mesmo percentual. Essa abordagem é fundamental para proteger o capital investido e garantir que o sistema opere em parâmetros controláveis, reduzindo a exposição a eventos extremos e à volatilidade excessiva.

3.4.4 Máximo Drawdown

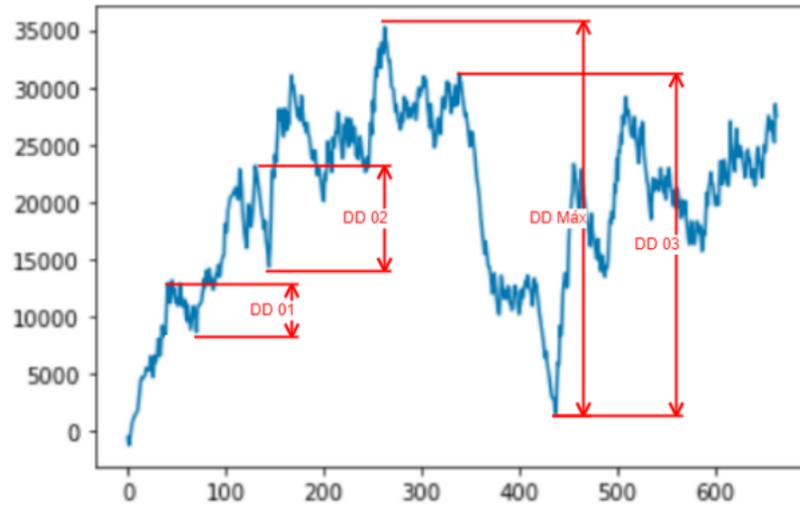
O drawdown representa a variação percentual negativa entre um pico e a mínima subsequente do capital ao longo do tempo, sendo uma métrica amplamente utilizada para avaliar o risco de estratégias de investimento. Dentro desse conceito, o Máximo Drawdown (MDD) refere-se à maior perda registrada de um pico até a menor mínima seguinte no período analisado.

O Máximo Drawdown é uma métrica fundamental para medir a estabilidade e robustez do sistema, pois indica a capacidade da estratégia de preservar o capital e evitar perdas significativas. Quanto menor o MDD, maior a resiliência do sistema a períodos de volatilidade e adversidade no mercado.

Além disso, essa métrica é especialmente relevante no contexto do Day Trade, onde oscilações abruptas podem comprometer rapidamente o desempenho do portfólio. Estratégias com

um MDD elevado tendem a expor o investidor a riscos mais acentuados, podendo dificultar a recuperação do capital perdido. Dessa forma, ao avaliar um sistema de negociação, é essencial considerar o Máximo Drawdown em conjunto com outras métricas, como o retorno médio e o índice de Sharpe, para uma análise mais abrangente da eficiência e do risco envolvido.

Figura 7 – Exemplo de Drawdowns.



Fonte: Albuquerque Filho (FILHO, 2021)

4 EXPERIMENTOS E RESULTADOS

Este capítulo tem como objetivo detalhar as configurações dos experimentos realizados, seguindo a metodologia apresentada no Capítulo 3, e discutir os resultados obtidos. Dado que a melhoria proposta neste trabalho impacta principalmente o sistema de Day Trade, a análise será focada nesse aspecto, explorando cada variação testada e os respectivos resultados para cada ação.

4.1 CONFIGURAÇÕES DOS EXPERIMENTOS

Para todos os experimentos realizados, mantivemos alguns parâmetros fixos, como a divisão do conjunto de dados em treino, validação e teste, conforme descrito na Seção 3.2. O conjunto de treino compreende 704 dias, enquanto os conjuntos de validação e teste possuem, cada um, 40 dias.

No Sistema de Predição, utilizamos um modelo composto por 300 árvores em todas as execuções, variando apenas os métodos de normalização, que serão detalhados na seção 4.3.

Para o Sistema de Day Trade, adotou-se um capital inicial de R\$500.000,00, um stop-loss de 2% e um desvio-padrão de 1% aplicado às previsões do modelo. O stop-loss foi utilizado para limitar perdas ao encerrar automaticamente operações que apresentassem variação negativa superior a 2% em relação ao preço de entrada — queda no caso de compras e alta no caso de vendas. Já o desvio-padrão funcionou como um fator de ajuste nos limites de entrada e saída, ampliando ou reduzindo levemente os preços previstos para compra e venda, com o objetivo de tornar o sistema mais sensível às variações de mercado.

4.2 FINAL SCORE

Para a avaliação do resultado final do sistema de Day Trading, foram consideradas três métricas principais:

1. **Montante Final:** Capital acumulado ao final dos 40 dias do conjunto de testes.
2. **Taxa de Sucesso:** Percentual de operações que resultaram em lucro.

3. **Máximo Drawdown:** Maior perda percentual entre um pico de capital e o menor valor subsequente.

Como o objetivo deste trabalho é destacar as configurações que melhor equilibram lucro e risco, foi adotado um sistema de pontuação ponderada entre as métricas. Para isso, atribuímos pesos distintos com base na relevância de cada indicador para fins de comparação entre os experimentos realizados. O **Montante Final**, por representar diretamente o ganho financeiro, recebeu o maior peso (6). A **Taxa de Sucesso** e o **Máximo Drawdown** receberam pesos 3 e 2, respectivamente, a fim de evitar cenários distorcidos, como estratégias que operam em poucos dias e apresentam 100% de acerto ou risco quase nulo, mas com baixo retorno financeiro.

É importante ressaltar que essa ponderação não representa uma regra universal para avaliação de estratégias de investimento. Trata-se apenas de um critério interno definido para este estudo, com o propósito de facilitar a análise comparativa entre os experimentos. Em contextos reais, os pesos ideais podem variar conforme o perfil de risco e os objetivos individuais de cada investidor.

$$\text{Score Final} = 6 \times \text{Montante Final} + 3 \times \text{Taxa de Sucesso} - 2 \times \text{Máximo Drawdown} \quad (4.1)$$

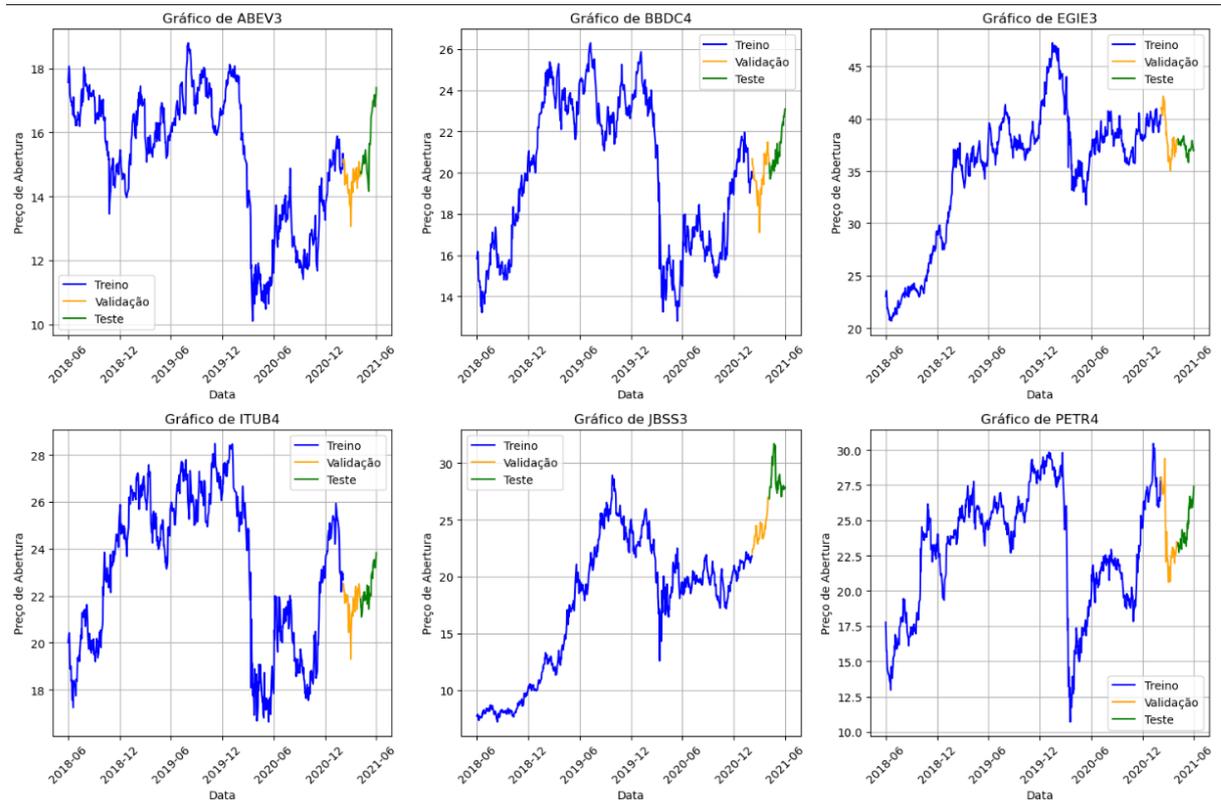
4.3 NORMALIZAÇÃO

Embora a normalização seja amplamente reconhecida como uma técnica eficaz em diversas aplicações, os resultados obtidos neste estudo indicam que, na maioria dos experimentos, seu impacto não foi positivo. Essa limitação pode ser atribuída à valorização significativa das ações ao longo dos anos (Figura 8), o que influencia diretamente a distribuição dos dados.

Ao normalizar os preços, especialmente em séries temporais financeiras, corre-se o risco de distorcer relações temporais importantes, comprometendo a capacidade do modelo de capturar padrões relevantes. No contexto deste trabalho, a normalização pode ter reduzido a sensibilidade do modelo a variações reais do mercado, dificultando a identificação de tendências e pontos de reversão nos preços das ações.

Dessa forma, os resultados sugerem que, para o problema específico abordado, a preservação dos valores originais pode ser mais benéfica do que a aplicação de técnicas tradicionais de normalização. Esse achado reforça a importância de avaliar criticamente o impacto da nor-

Figura 8 – Preço de abertura das ações ao longo do tempo.



Fonte: Produzido pelo próprio Autor.

malização em modelos financeiros e considerar abordagens alternativas, como transformações específicas que preservem a escala relativa dos preços sem comprometer a estrutura temporal dos dados.

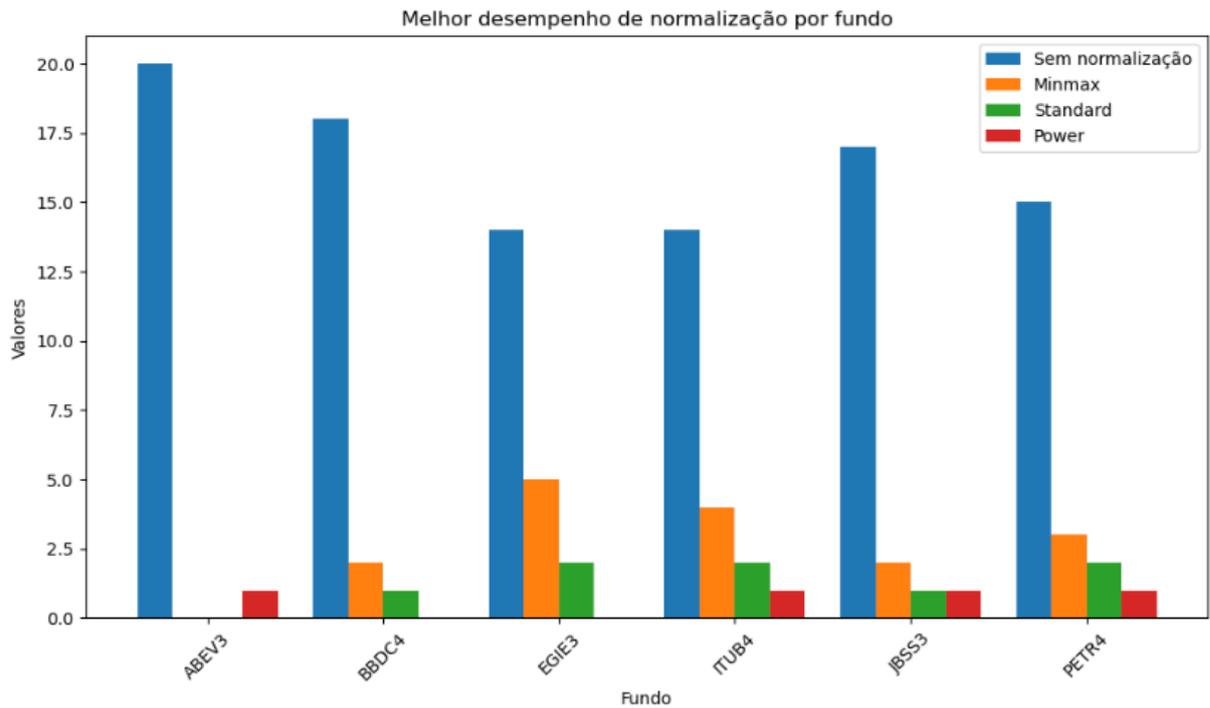
Para avaliar o impacto dos diferentes tipos de normalização em cada ação, consideramos o resultado com a melhor final score dentro de cada intervalo de incerteza. Com isso, elaboramos um ranking que permite comparar o desempenho de cada abordagem de normalização de forma estruturada, evidenciado na Figura 9.

No entanto, observamos resultados excepcionais em ações como ITUB4, JBSS3 e PETR4, que superaram aquelas que não utilizaram normalização, conforme ilustrado na tabela 5.

Tabela 5 – Melhores resultados com e sem normalização.

Ação	Normalização	Montante Final	Taxa de Sucesso (%)	Drawdown (%)	Número de dias Operados	Final Score
ITUB4	Standard	519.238,00	50	3,89	15	0,76
	Não Utilizada	499.275,00	53,57	4,00	16	0,75
JBSS3	MinMax	562.809,00	60	2,68	9	0,85
	Não Utilizada	564.322,99	57,89	4,80	15	0,84
PETR4	Standard	542.533,00	72,72	4,18	19	0,86
	Não Utilizada	522.809,00	60,76	7,03	33	0,79

Figura 9 – Frequência das normalizações que obtiveram o melhor Final Score por ação.



Fonte: Produzido pelo próprio Autor.

4.4 INTERVALO DE INCERTEZA

Nesta seção, avaliamos os impactos da escolha do intervalo de incerteza na performance das operações, comparando os resultados para diferentes valores desse parâmetro. Para garantir uma análise consistente, utilizamos experimentos sem normalização, conforme discutido na Seção 4.3.

Tabela 6 – ABEV3

Incerteza	Acurácia	MDD	Dias Operados
5.00%	100.00%	0.00%	1
15.00%	69.57%	2.4585%	9
25.00%	65.85%	2.7835%	17
35.00%	62.07%	6.2958%	27
45.00%	59.15%	6.4722%	35
55.00%	59.15%	6.4722%	35
65.00%	60.53%	6.4710%	39
75.00%	60.53%	6.4710%	39
85.00%	60.53%	6.4710%	39
95.00%	60.26%	6.4670%	40

Tabela 7 – BBDC4

Incerteza	Acurácia	MDD	Dias Operados
5.00%	55.56%	1.9056%	5
15.00%	52.94%	2.6473%	20
25.00%	55.17%	2.4064%	34
35.00%	53.73%	4.7574%	40
45.00%	53.73%	4.7574%	40
55.00%	53.73%	4.7574%	40
65.00%	53.73%	4.7574%	40
75.00%	53.73%	4.7574%	40
85.00%	53.73%	4.7574%	40
95.00%	53.73%	4.7574%	40

Tabela 8 – EGIE3

Incerteza	Acurácia	MDD	Dias Operados
5.00%	76.00%	0.6744%	4
15.00%	75.86%	1.8302%	16
25.00%	77.57%	3.1077%	33
35.00%	77.69%	3.1050%	39
45.00%	78.05%	3.1062%	40
55.00%	78.05%	3.1062%	40
65.00%	78.05%	3.1062%	40
75.00%	78.05%	3.1062%	40
85.00%	78.05%	3.1062%	40
95.00%	78.05%	3.1062%	40

Tabela 9 – ITUB4

Incerteza	Acurácia	MDD	Dias Operados
5.00%	37.50%	4.3379%	6
15.00%	53.57%	4.0018%	16
25.00%	48.57%	4.6100%	22
35.00%	50.00%	5.3092%	24
45.00%	47.73%	8.4979%	29
55.00%	49.02%	8.9243%	35
65.00%	50.94%	8.1393%	37
75.00%	50.91%	7.7313%	39
85.00%	51.79%	7.4673%	40
95.00%	51.79%	7.4673%	40

Tabela 10 – JBSS3

Incerteza	Acurácia	MDD	Dias Operados
5.00%	0.00%	0.0000%	0
15.00%	0.00%	0.0000%	0
25.00%	0.00%	0.0000%	0
35.00%	0.00%	0.0000%	0
45.00%	0.00%	0.0000%	0
55.00%	57.89%	4.7980%	15
65.00%	52.00%	5.3644%	20
75.00%	47.06%	7.2278%	27
85.00%	42.86%	9.4208%	33
95.00%	47.92%	9.4208%	37

Tabela 11 – PETR4

Incerteza	Taxa de Sucesso	MDD	Dias Operados
5.00%	63.64%	6.7612%	15
15.00%	60.76%	7.0319%	33
25.00%	56.82%	10.0402%	39
35.00%	57.30%	9.3157%	40
45.00%	57.30%	9.3157%	40
55.00%	57.30%	9.3157%	40
65.00%	57.30%	9.3157%	40
75.00%	57.30%	9.3157%	40
85.00%	57.30%	9.3157%	40
95.00%	57.30%	9.3157%	40

Os resultados apresentados nas tabelas acima indicam que o máximo drawdown tende a diminuir à medida que a incerteza restringe as operações. Isso sugere que valores mais conservadores de incerteza podem reduzir a exposição a quedas expressivas no capital.

Além disso, observa-se uma correlação entre a taxa de sucesso de operações e o intervalo de incerteza na maioria dos casos. No entanto, essa relação não é universal, uma vez que alguns intervalos específicos apresentam variações inesperadas na taxa de sucesso. Esses desvios indicam que o impacto da incerteza pode ser influenciado por outros fatores, como a volatilidade do ativo e a dinâmica do mercado no período analisado.

Dessa forma, a escolha do intervalo de incerteza deve buscar um equilíbrio entre a redução de risco (drawdown), a preservação da performance operacional (taxa de sucesso e número de

operações) e a maximização dos ganhos de capital.

4.5 RESULTADOS FINAIS

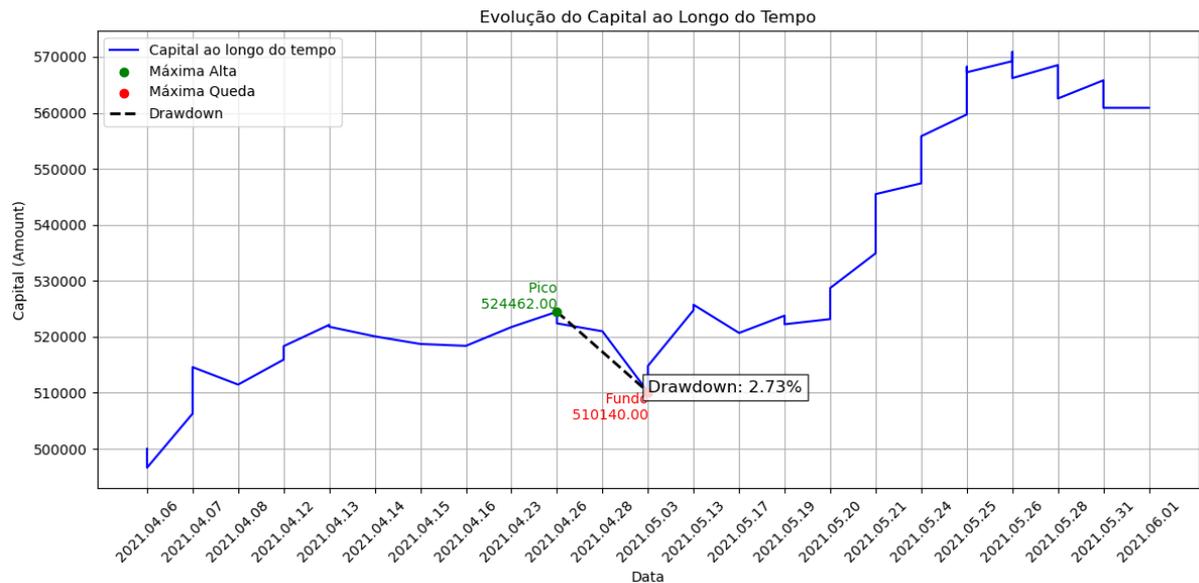
Para estabelecer uma conclusão robusta, compararemos o melhor resultado obtido para todas as ações no trabalho original, sem a aplicação da incerteza, com um dos melhores resultados alcançados ao introduzi-la. A análise abrangerá tanto cenários com e sem normalização, além de considerar a maximização do número de dias com operações ativas. Dessa forma, será possível avaliar o impacto da incerteza na performance do modelo, levando em conta fatores como retorno final, redução de risco (máximo drawdown) e taxa de sucesso das operações.

1. ABEV3

Tabela 12 – Resultados ABEV3.

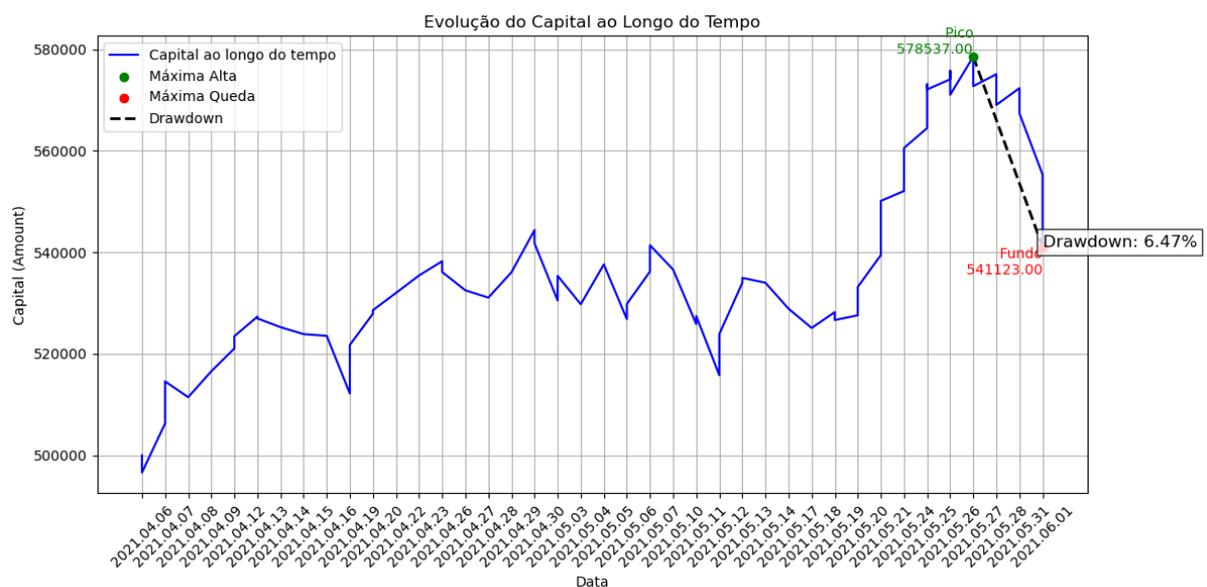
Incerteza Threshold	Capital Final	Taxa de Sucesso (%)	TradedDays	Normalização	MDD (%)	FinalScore
Intervalo 30%	R\$560.897	66,0	22	None	2,73	0,86
Não utilizada	R\$541.123	60,26	40	None	6,46	0,81

Figura 10 – Resultado ABEV3 usando intervalo de incerteza.



Fonte: Produzido pelo próprio Autor.

Figura 11 – Melhor resultado ABEV3 sem utilizar incerteza.



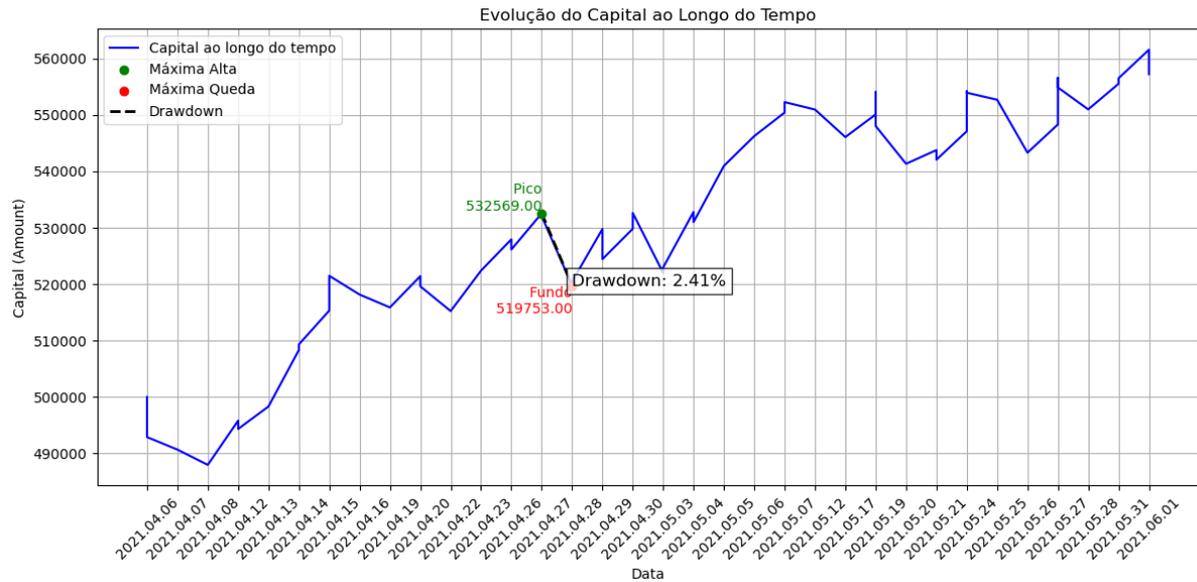
Fonte: Produzido pelo próprio Autor.

2. BBDC4

Tabela 13 – Resultados BBDC4.

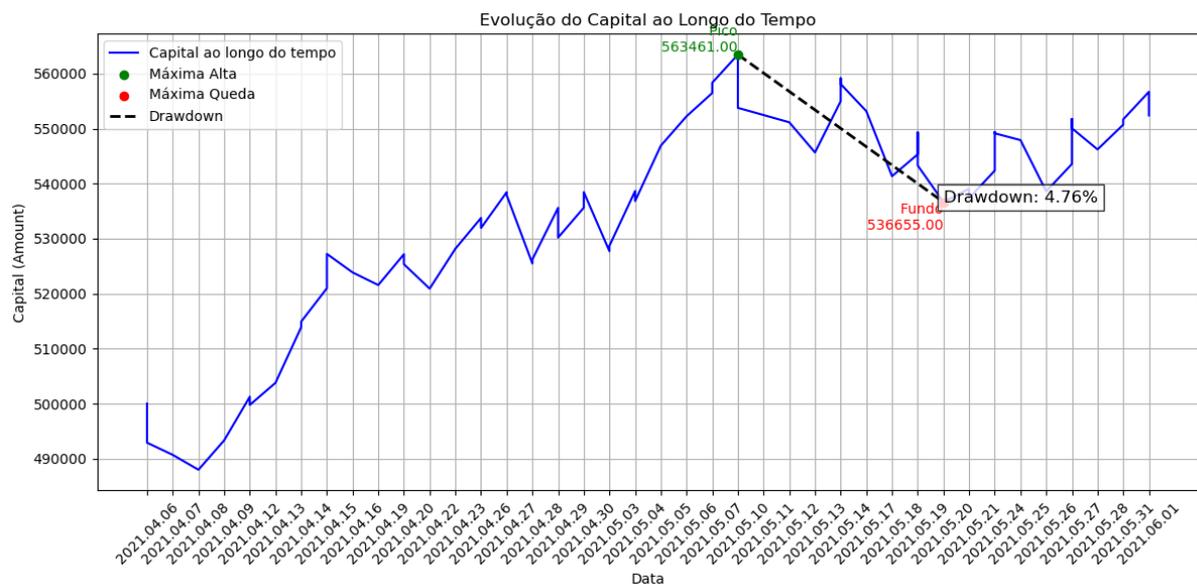
Incerteza Threshold	Capital Final	Taxa de Sucesso (%)	TradedDays	Normalização	MDD (%)	FinalScore
Intervalo 25%	R\$557.222	55,17	34	None	2,41	0,82
Não utilizada	R\$552.403	53,73	40	None	4,76	0,81

Figura 12 – Resultado BBDC4 usando intervalo de incerteza.



Fonte: Produzido pelo próprio Autor.

Figura 13 – Melhor resultado BBDC4 sem utilizar incerteza.



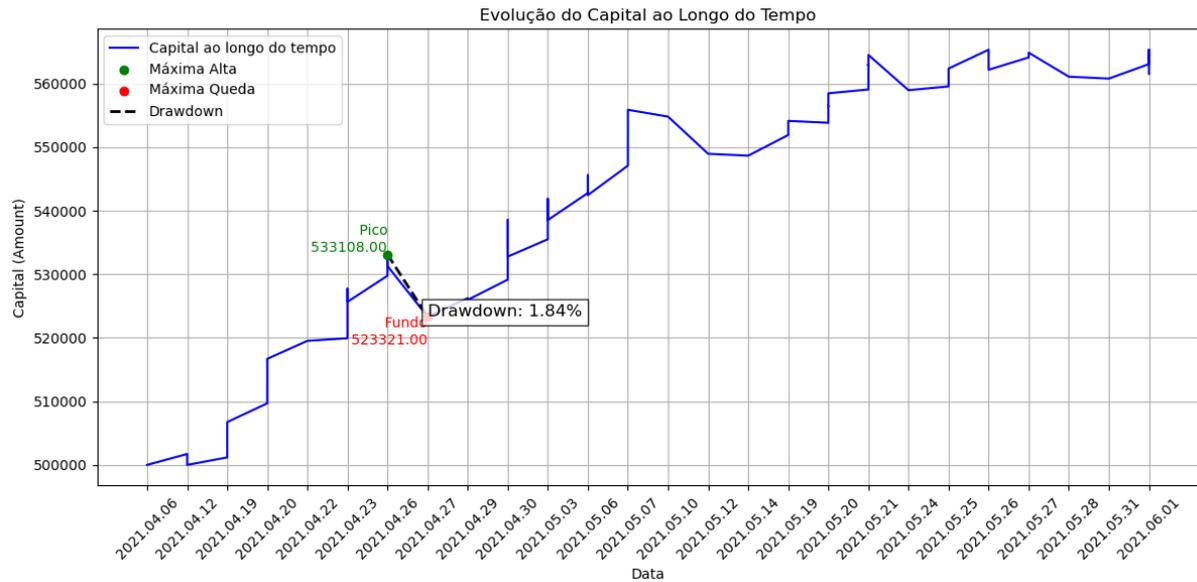
Fonte: Produzido pelo próprio Autor.

3. EGIE3

Tabela 14 – Resultados EGIE3.

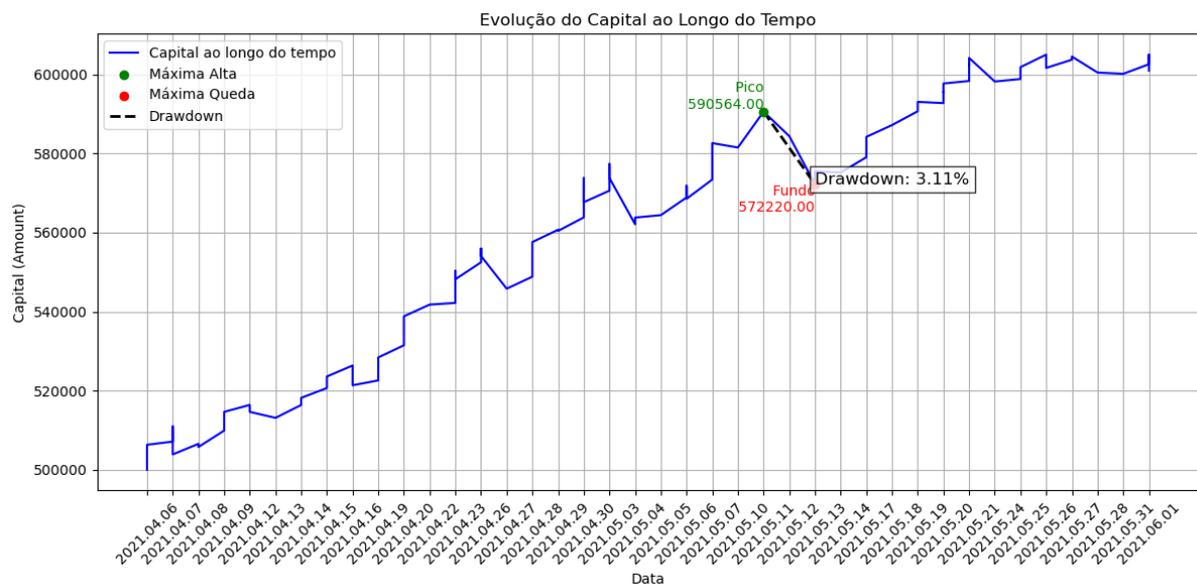
Incerteza Threshold	Capital Final	Taxa de Sucesso (%)	TradedDays	Normalização	MDD (%)	FinalScore
Intervalo 20%	R\$561.544	76,14	25	None	1,84	0,90
Não utilizada	R\$600.914	78,05	40	None	3,11	0,95

Figura 14 – Resultado EGIE3 usando intervalo de incerteza.



Fonte: Produzido pelo próprio Autor.

Figura 15 – Melhor resultado EGIE3 sem utilizar incerteza.



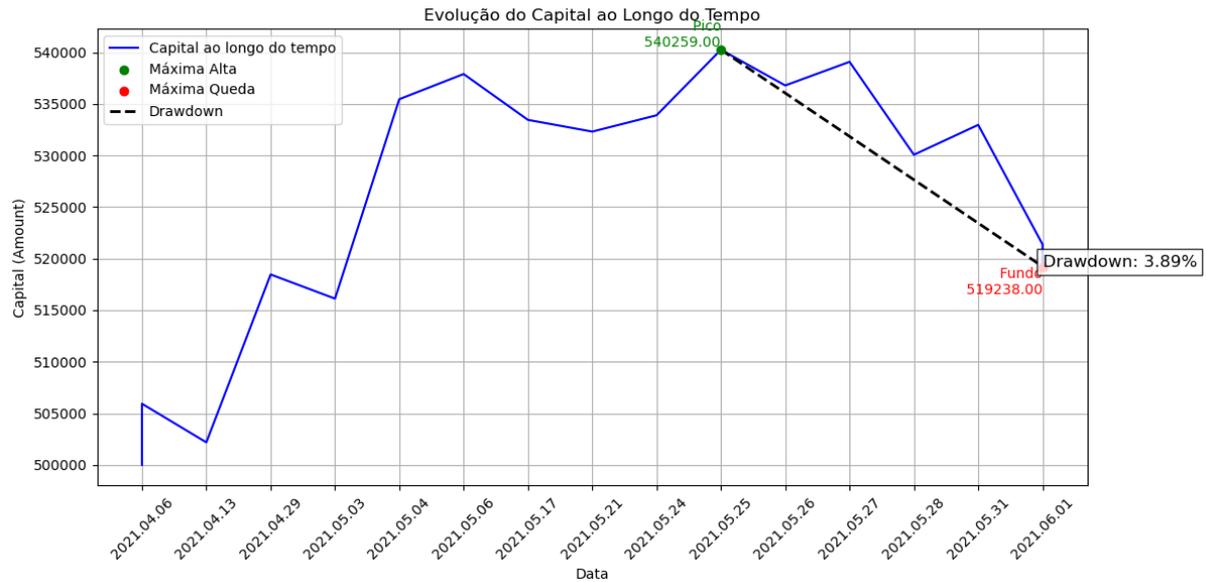
Fonte: Produzido pelo próprio Autor.

4. ITUB4

Tabela 15 – Resultados ITUB4.

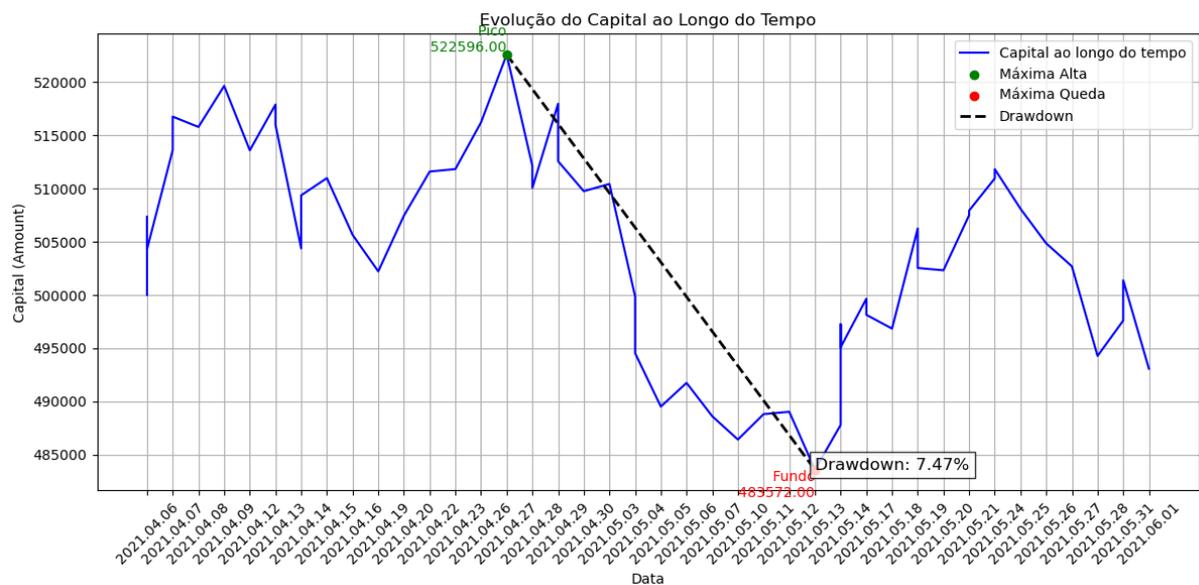
Incerteza Threshold	Capital Final	Taxa de Sucesso (%)	TradedDays	Normalização	MDD (%)	FinalScore
Intervalo 10%	R\$519.238	50,0	15	Standard	3,89	0,76
Não utilizada	R\$493.066	51,79	40	None	7,47	0,73

Figura 16 – Resultado ITUB4 usando intervalo de incerteza.



Fonte: Produzido pelo próprio Autor.

Figura 17 – Melhor resultado ITUB4 sem utilizar incerteza.



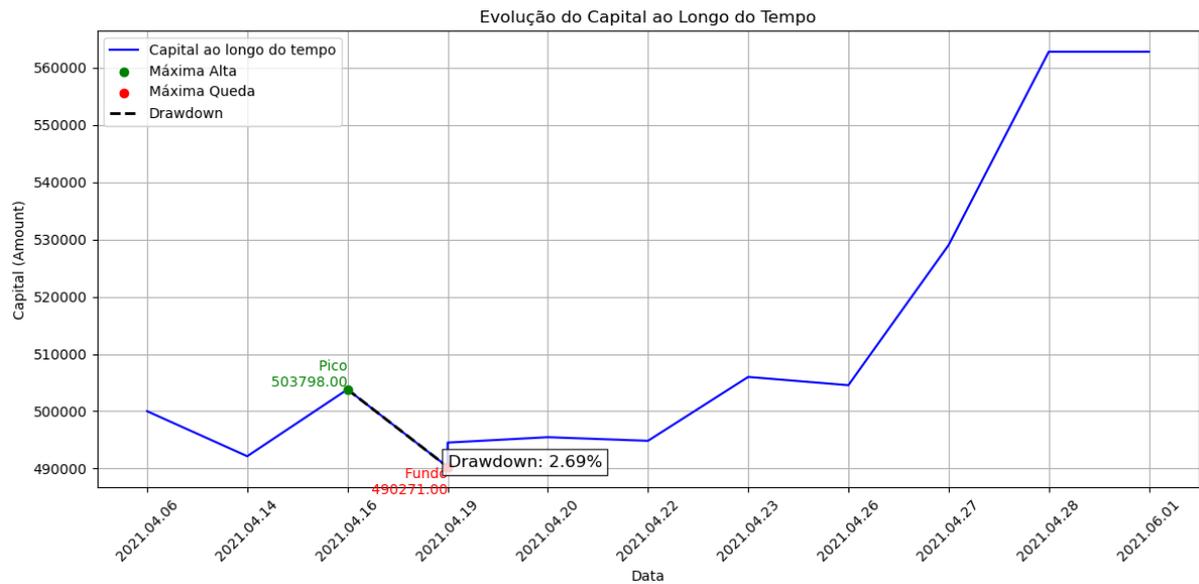
Fonte: Produzido pelo próprio Autor.

5. JBSS3

Tabela 16 – Resultados JBSS3.

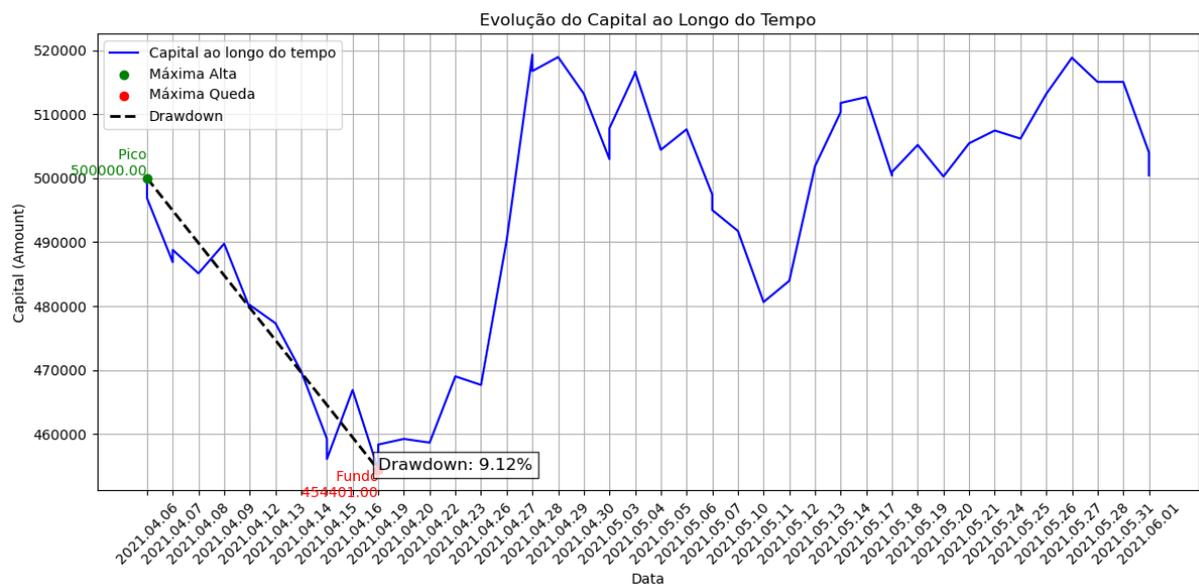
Incerteza Threshold	Capital Final	Taxa de Sucesso (%)	TradedDays	Normalização	MDD (%)	FinalScore
Intervalo 45%	R\$562.809	60,0	9	Minmax	2,69	0,85
Não utilizada	R\$500.445	49,02	40	None	9,12	0,73

Figura 18 – Resultado JBSS3 usando intervalo de incerteza.



Fonte: Produzido pelo próprio Autor.

Figura 19 – Melhor resultado JBSS3 sem utilizar incerteza.



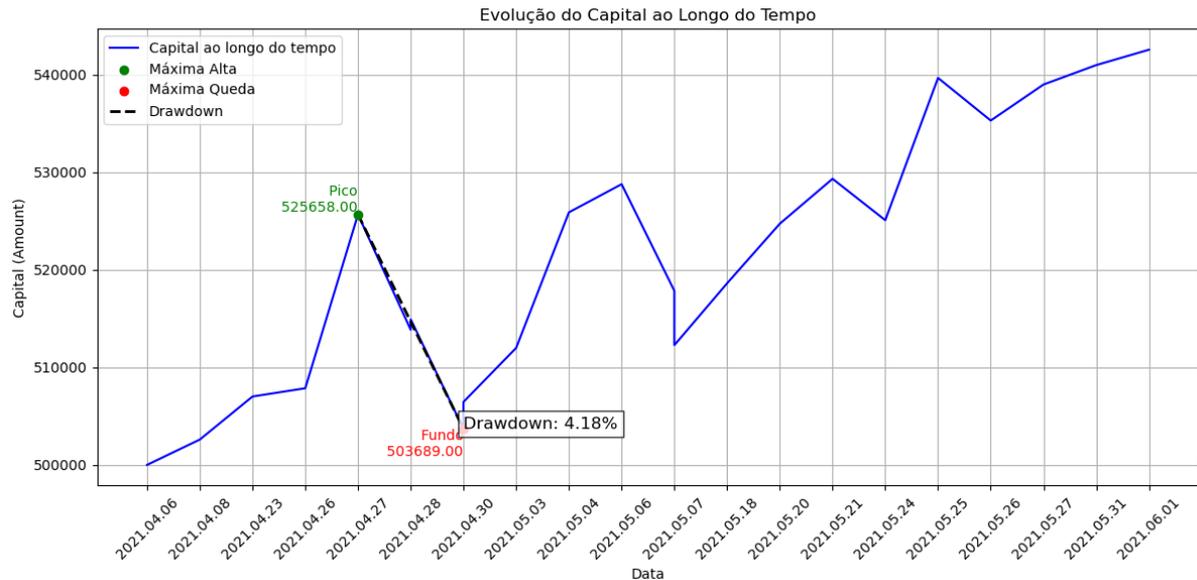
Fonte: Produzido pelo próprio Autor.

6. PETR4

Tabela 17 – Resultados PETR4.

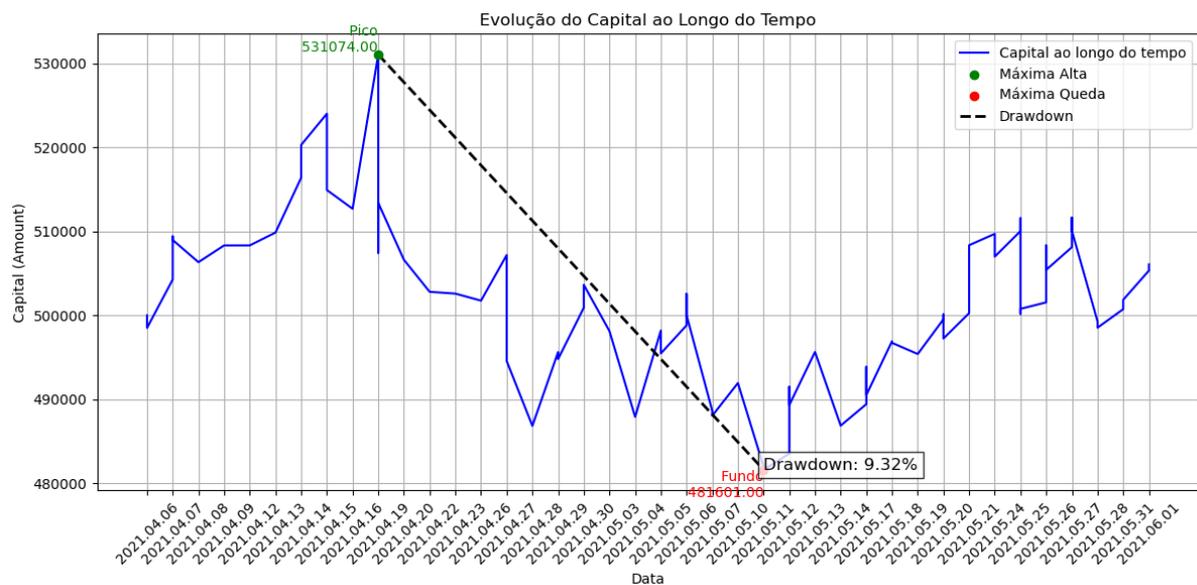
Incerteza Threshold	Capital Final	Taxa de Sucesso (%)	TradedDays	Normalização	MDD (%)	FinalScore
Intervalo 5%	R\$542.533	72,73	19	Standard	4,18	0,86
Não utilizada	R\$506.076	57,3	40	None	9,32	0,76

Figura 20 – Resultado PETR4 usando intervalo de incerteza.



Fonte: Produzido pelo próprio Autor.

Figura 21 – Melhor resultado PETR4 sem utilizar incerteza.



Fonte: Produzido pelo próprio Autor.

Ao analisar os resultados, observamos uma redução média de 54% no máximo drawdown quando utilizamos a incerteza da forma proposta nesse trabalho, indicando uma significativa mitigação do risco associado às operações. Além disso, houve um aumento médio de 3,81% no ganho de capital, com o melhor cenário apresentando um acréscimo de até 12% (Tabela 16). Esses resultados evidenciam que a adoção da incerteza pode contribuir tanto para a preservação do capital, reduzindo a exposição a perdas expressivas, quanto para a melhoria do desempenho financeiro, potencializando os retornos obtidos.

5 TRABALHOS FUTUROS

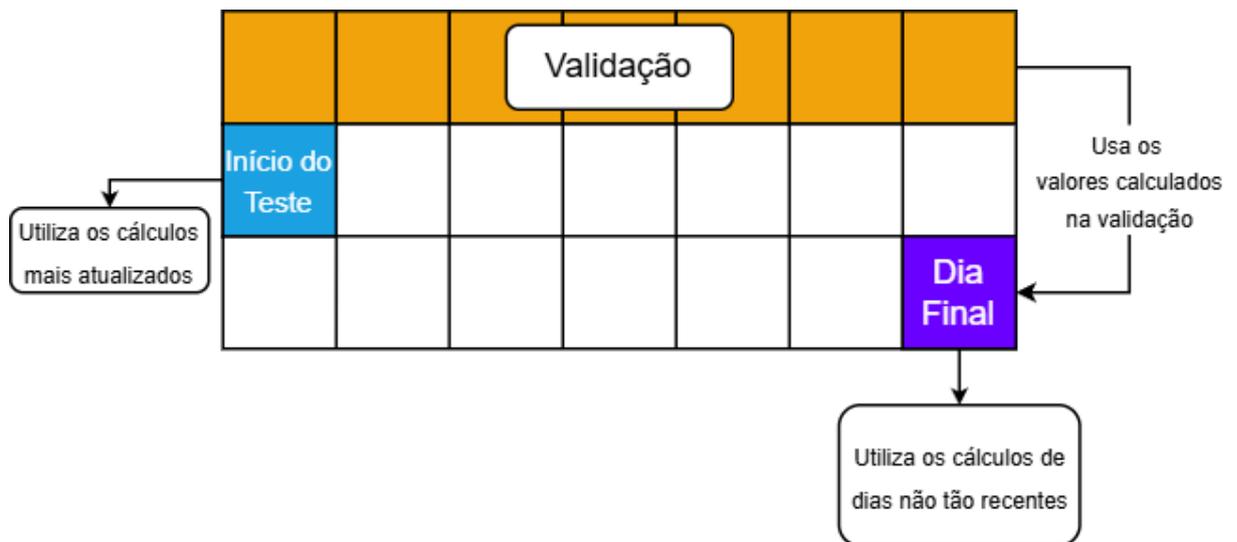
Os resultados obtidos nesta pesquisa demonstraram que o uso da incerteza é eficaz na maximização dos lucros e na mitigação de riscos. No entanto, é crucial destacar a amplitude e as limitações dos experimentos realizados.

Para futuras pesquisas, seria interessante expandir o uso da incerteza para outros ativos e ações, além de testar o sistema com diferentes periodicidades de dados, além do intervalo de 15 minutos utilizado nesta análise. Isso permitiria uma compreensão mais abrangente do impacto da incerteza em mercados com diferentes características e comportamentos.

Outro ponto importante a ser considerado é que o sistema não foi testado para simular uma taxa de retorno anualizada, principalmente por limitações de tempo e, mais crucialmente, por questões lógicas relacionadas ao cálculo da incerteza. Como detalhado na Seção 3.2, os intervalos de incerteza foram calculados a partir de apenas 40 dias do conjunto de validação. Como discutido na Seção 3.3.1.3, a incerteza depende de previsões para ser calculada. Dada a volatilidade do mercado, surge uma dúvida sobre a eficácia de utilizar um intervalo tão curto para a aplicação em períodos mais longos, onde as variações de preço são mais significativas.

Nesse contexto, seria interessante desenvolver um método para retreinamento e recalibração do sistema preditivo em intervalos regulares — como diário, semanal ou mensal. Essa abordagem permitiria que os intervalos de incerteza fossem atualizados com base nas novas previsões do modelo, tornando sua aplicação mais precisa no sistema de day trading. Assim, os intervalos percentuais utilizados permaneceriam como referência, porém ajustados de forma dinâmica, acompanhando as oscilações do mercado e promovendo uma gestão de risco mais responsiva e adaptável. Como ilustrado na Figura 22, a falta de recalibração com dados recentes pode comprometer a eficácia dos intervalos de incerteza ao longo do tempo.

Figura 22 – Falha no uso da incerteza.



Fonte: Produzido pelo próprio Autor.

REFERÊNCIAS

- CHANDE, T. S. *Beyond Technical Analysis: How to Develop and Implement a Winning Trading System*. 2. ed. New York: John Wiley & Sons, 2001.
- COULSTON, J. W.; BLINN, C. E.; THOMAS, V. A.; WYNNE, R. H. Approximating prediction uncertainty for random forest regression models. *Photogrammetric Engineering and Remote Sensing*, American Society for Photogrammetry and Remote Sensing, v. 82, n. 3, p. 189–197, mar. 2016. ISSN 0099-1112. Disponível em: <<http://dx.doi.org/10.14358/PERS.82.3.189>>.
- DUTSCHMANN, T. M.; BAUMANN, K. Evaluating high-variance leaves as uncertainty measure for random forest regression. *Molecules*, v. 26, n. 21, p. 6514, Oct 2021.
- DÍAZ, J. D.; HANSEN, E.; CABRERA, G. Machine-learning stock market volatility: Predictability, drivers, and economic value. *International Review of Financial Analysis*, v. 94, n. C, 2024. Disponível em: <<https://ideas.repec.org/a/eee/finana/v94y2024ics1057521924002187.html>>.
- FILHO, A. C. d. A. *Avaliação do método de random forest para indicar o melhor momento de compra e venda de ações no mercado financeiro*. Trabalho de Conclusão de Curso — Universidade Federal de Pernambuco (UFPE), Centro de Informática, 2021.
- JUNIOR, L. A. d. A. *Sistema automático para operações day trade baseado em previsão de valores máximos e mínimos com random forest*. Trabalho de Conclusão de Curso — Universidade Federal de Pernambuco (UFPE), Centro de Informática, 2022.
- MARTINEZ, L. C.; HORA, D. N. da; PALOTTI, J. R. de M.; MEIRA, W.; PAPPA, G. L. From an artificial neural network to a stock market day-trading system: A case study on the bmf bovespa. In: *2009 International Joint Conference on Neural Networks*. [S.l.: s.n.], 2009. p. 2006–2013.
- VIJH, M.; CHANDOLA, D.; TIKKIWAL, V.; KUMAR, A. Stock closing price prediction using machine learning techniques. *Procedia Computer Science*, v. 167, p. 599–606, 01 2020.