

Determinantes do desempenho de modelos baseados em algoritmos de inteligência artificial em finanças corporativas: estudo nos mercados de capitais do G20 no século XXI

Autores:

Ewerton Alex Avelar

ewertonalexavelar@gmail.com

Joyce Mariella Medeiros Cavalcanti

joyce.cavalcanti@unp.br

Ricardo Vinícius Dias Jordão

jordaoconsultor@yahoo.com.br

Terence Machado Boina

tmboina@gmail.com

Beatriz Mitsuyo Fuzikawa Bicalho Cotta

beatriz.bcotta@hotmail.com

Resumo: A pesquisa apresentada neste artigo analisou os determinantes do desempenho dos modelos baseados em algoritmos de inteligência artificial (IA) para previsão de preços dos principais índices das bolsas de valores de países membros do G20 ao longo do século XXI. Para isso, desenvolveram-se: (1) modelos baseados em algoritmos de IA (*Naive Bayes*, *Random Forest*, e *k-Nearest Neighbors*), cujos desempenhos foram mensurados por acurácia, precisão, *Recall* e F1-Score; e (2) modelos de regressão para estimar a relação entre o desempenho dos modelos desenvolvidos na primeira etapa e variáveis determinantes. Os resultados evidenciaram que: (a) o desempenho dos modelos baseados em algoritmos de IA foi superior ao esperado pelo mercado; (b) a capacidade de explicação variou consistentemente de acordo com a métrica empregada para mensurar desempenho; (c) o desempenho dos modelos preditivos diminuiu quando houve um número maior de retornos negativos (quedas); e (d) as demais variáveis determinantes apresentaram resultados heterogêneos. Dentre as contribuições, destacaram-se: (i) ineditismo em analisar variáveis determinantes do desempenho de modelos baseados em algoritmos de IA; (ii) evidências robustas sobre desempenho consistente e sistematicamente superior desses modelos baseados em relação ao mercado; e (iii) poder preditivo oferecido por algoritmos de IA aos indicadores técnicos para previsão de preços.

Palavras-chave: Algoritmos de inteligência artificial (IA); Previsão de preços; Determinantes de desempenho; Grupo dos Vinte (G20).

1. Introdução

A capacidade de prever os preços dos ativos nos mercados financeiros possibilita vários ganhos potenciais aos agentes econômicos (Beniwal et al., 2024). Contudo, a previsão de preços nesse ambiente constitui um desafio relevante para agentes de mercado e para a própria Academia, dada a multiplicidade de fatores que afetam esses valores, como apontado por Srivastava et al. (2023) e Shaban et al. (2024). Nesse cenário, nos últimos anos, modelos apoiados em algoritmos de inteligência artificial (IA) têm se consolidado como opções de destaque para previsão de preços no mercado financeiro (Ferreira et al., 2021). Ainda assim, nesse contexto, deve-se enfatizar que a hipótese de mercados eficientes (HME) estabelece que não seria possível obter retornos anormais de forma consistente ao se tentar prever preços de ativos. Segundo Hung et al. (2024), sob essa hipótese, o preço dos ativos refletiria integralmente todas as informações relevantes disponíveis no mercado.

Por outro lado, contrariando a forma fraca da HME, diversos trabalhos têm mostrado que algoritmos de IA aplicados a previsões alcançam desempenhos superiores à média esperada pelo mercado (e.g., Shynkevich et al., 2017; Cao et al., 2019; Mndawe et al., 2022; Srivastava et al., 2023; Hung et al., 2024), estimulando um debate crescente sobre o tema. Muitos desses estudos utilizam indicadores técnicos como variáveis de entrada (*inputs*) dos modelos (e.g., Hsu et al., 2016; Chernavin, 2019; Ampomah et al., 2021; Khoa & Huynh, 2021; Avelar & Jordão, 2024). Beniwal et al. (2024) observaram que tais indicadores se baseiam em informações históricas para projetar preços futuros, desafiando diretamente os pressupostos da HME.

Em geral, tais estudos empregam métricas típicas de aprendizado de máquina (*machine learning*) para mensurar o desempenho dos modelos de IA e justificar sua utilidade no mercado financeiro (e.g., Chernavin, 2019; Wu et al., 2020; Khoa & Huynh, 2021; Avelar & Jordão, 2024; Sagaceta-Mejía et al., 2024), tais como acurácia, precisão, *Recall* e F1-Score. Tais métricas representam diferentes formas de se mensurar o desempenho de um modelo de IA (Burger, 2018; Lantz, 2019; Nwanganga & Chapple, 2020) e têm apresentado resultados bastante divergentes entre si e entre os modelos desenvolvidos em diferentes estudos, tais como: Patel et al. (2015), Nabipour et al. (2017), Wang et al. (2018), Ampomah et al. (2021), Wu et al. (2020), Awan et al. (2021) e Avelar e Jordão (2024). Contudo, não foram identificados estudos que analisem as variáveis que possam explicar as variações de desempenho dos modelos de IA observadas nesses estudos sobre previsão no mercado financeiro, tornando-se essencial aprofundar a análise desse tema e explorar essa relevante lacuna de pesquisa.

Nesse sentido, a pesquisa cujos resultados são apresentados neste artigo visou analisar os determinantes do desempenho dos modelos baseados em algoritmos de IA, com base em uma análise dos principais índices das bolsas de valores de países do Grupo dos Vinte (G20) ao longo do século XXI. Para tanto, foram propostos e cumpridos os seguintes objetivos específicos: (a) identificar variáveis com o potencial de influenciar o desempenho de algoritmos de IA; (b) verificar a relação entre tais variáveis e o desempenho desses algoritmos; (c) comparar o papel das diferentes variáveis selecionadas sob as perspectivas das diferentes métricas de desempenho; e (d) discutir possíveis oportunidades e limitações do emprego de algoritmos de IA para a previsão de preços de ativos na pesquisa em finanças corporativas.

Assim, a pesquisa ora apresentada inova ao focar os determinantes do desempenho de modelos baseados em algoritmos de IA, cada vez mais empregados na Academia e no Mercado para previsão de preços. Dessa forma, contribui-se com a área de Finanças Corporativas ao gerar novas evidências e discutir sobre as implicações práticas e teóricas do

desempenho desses modelos, considerando a ampla adoção dessa tecnologia no mercado financeiro e o debate persistente sobre a HME.

2. Fundamentação teórica

A maior parte dos estudos que empregaram IA para previsão de preços no mercado recorreram a algoritmos de aprendizado supervisionado, usando indicadores técnicos como *inputs* dos modelos (e.g., Patel et al., 2015; Hsu et al., 2016; Nabipour et al., 2017; Ampomah et al., 2021; Singh & Khushi, 2021; Shaban et al., 2024). Dentre os indicadores técnicos empregados em estudos como os de Patel et al. (2015), Nabipour et al. (2020), Chernavin (2019), Mndawe et al. (2022) e Sagaceta-Mejía et al. (2024), podem ser destacados: *Simple Moving Average* (SMA), *Weighted Moving Average* (WMA), *Moving Average Convergence/Divergence* (MACD), *Momentum Stochastic K%* (STCK), *Momentum Stochastic D%* (STCD), *Relative Strength Index* (RSI), *William's Percent R* (WPR) e *Commodity Channel Index* (CCI).

Destaca-se, ainda, que têm sido empregados diferentes algoritmos de IA para fins de previsão de preços de ações no mercado de capitais, tais como Naive Bayes (NB), *K-nearest Neighbors* (KNN) e *Random forest* (RF) (Shynkevich et al., 2017; Weng et al., 2017; Avelar & Jordão, 2024; Shaban et al., 2024). Usando modelos baseados em tais algoritmos e empregando os indicadores técnicos supracitados, diversos desses estudos vêm reportando alto desempenho de capacidade preditiva ao longo do século XXI, potencializando retornos anormais, como observado por Qian e Rasheed (2007), Shynkevich et al. (2017), Cao et al. (2019), Chernavin (2019), Mndawe et al. (2022) e Sagaceta-Mejía et al. (2024), em desacordo com a HME. Contudo, observam-se divergências relevantes entre o desempenho reportado por tais estudos, indo desde métricas com valores pouco superiores a 50%-60% (e.g., Shynkevich et al., 2017; Wang et al., 2018; Wu et al., 2020) a outros valores superiores a 80%-90% (e.g., Qian & Rasheed, 2007; Patel et al., 2015; Ampomah et al., 2021; Awan et al., 2021; Avelar & Jordão, 2024).

Ressalta-se que tais estudos evidenciaram o desempenho de seus modelos a partir de diferentes métricas baseadas na literatura de aprendizado de máquina (Cf. Burger, 2018; Lantz, 2019; Nwanganga & Chapple, 2020). Para fins de classificação como a previsão de movimento de alta ou baixa do preço de um ativo (objetivo dos estudos citados acima, por exemplo), são usualmente empregadas as seguintes métricas: (a) Acurácia – proporção de observações que foram corretamente classificadas pelo preditor (Equação 1); (b) Precisão (*Precision*) – proporção de observações preditas na classe positiva e que realmente pertencem a esta classe estimada (Equação 2); (c) Revocação ou sensibilidade (*Recall*) – proporção de elementos da classe positiva que foram previstos como positivos (Equação 3); e (d) F1-Score – média harmônica entre o *Precision* e o *Recall* (Equação 4). Salienta-se que, em todas as equações: ‘VP’ equivale a “verdadeiros positivos”; ‘VN’ equivale a “verdadeiros negativos”; ‘FP’ equivale a “falsos positivos”; e ‘FN’ equivale a “falsos negativos”.

$$\text{Acurácia} = \frac{VP + VN}{VP + FP + VN + FN} \quad (1)$$

$$\text{Precision} = \frac{VP}{VP + FP} \quad (2)$$

$$\text{Recall} = \frac{VP}{VP + FN} \quad (3)$$

$$\text{F1 - Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Apesar de tais métricas supracitadas serem amplamente empregadas em estudos sobre previsão de preços de ativos envolvendo modelos de IA (como os estudos previamente citados), verifica-se que cada medida enfatiza (penaliza) os erros de classificação de forma distinta. Por exemplo, enquanto o *precision* foca na redução de FP, o *Recall* se direciona à redução dos FN. Em contrapartida, embora a acurácia trate os FN e os FP com o mesmo peso e seja uma medida intuitiva, seu resultado pode ser enganoso em dados desbalanceados. Tais variações de enfoque das medidas podem ser responsáveis por resultados bastante discrepantes no desempenho reportado por diferentes estudos entre as métricas, tais como: Patel et al. (2015), Nabipour et al. (2017), Wang et al. (2018), Wu et al. (2020) e Avelar e Jordão (2024). Dessa forma, uma vez que essas métricas são essenciais para desenvolvimento dos modelos de classificação e amplamente empregados nos estudos para previsão de movimento dos preços de ativos, formularam-se as hipóteses 1a e 1b:

H_{1a} – O maior número de quedas no período afeta negativa e significativamente o desempenho dos modelos baseados em algoritmos de IA.

H_{1b} – O maior número de altas no período afeta positiva e significativamente o desempenho dos modelos baseados em algoritmos de IA.

A literatura de modelos de IA destaca a importância dos dados usados como variáveis de entrada para o desempenho dos modelos (Anysz et al., 2016; Ali, 2022; Imani et al., 2024). Burger (2018) chegou a afirmar que a qualidade dos dados empregados para treinar os modelos é mais importante que os algoritmos de IA empregados para gera-los. Um aspecto bastante destacado na literatura se refere ao nível de variabilidade (dispersão) dos dados (Anysz et al., 2016; Novak & Velušček, 2016; Lantz, 2019; Nwanganga & Chapple, 2020) que pode prejudicar o desempenho dos modelos. Assim, desenvolveu-se a Hipótese 2 (H2):

H₂ – A maior dispersão dos dados no período afeta negativa e significativamente o desempenho dos modelos baseados em algoritmos de IA.

Salienta-se que diversos estudos que empregaram modelos baseados em algoritmos de IA para prever preços dos ativos consideraram empresas de diferentes países em sua amostra, tais como: Hsu et al. (2016), Mohanty et al. (2021), Avelar e Jordão (2024) e Gupta e Kumar (2024). Destaca-se, porém, que, em estudos que envolvem diferentes nações, é importante considerar aspectos da governança de cada uma, devido às peculiaridades inerentes às mesmas (Kaufmann et al., 2011). A governança pode ser definida pelo como as tradições e instituições pelas quais a autoridade em um país é exercida (The World Bank, 2024) e pode ser mensurada por meio do *Worldwide Governance Indicators* conforme Kaufmann e Kraay (2023). Diversos estudos demonstraram que tais tradições e instituições têm um papel relevante no valor das empresas, mensurado pelos preços dos ativos, em diferentes mercados de capitais ao redor do mundo, tais como: Saona e Martín (2018), Ojeka et al. (2019), Thakur et al. (2021) e Bagh et al. (2023). Desse modo, foi proposta a Hipótese 3:

H₃ – O maior nível de qualidade institucional afeta positiva e significativamente o desempenho dos modelos baseados em algoritmos de IA.

Por sua vez, a liberdade econômica foca no impacto da liberdade e dos mercados livres em todo o mundo sendo também importante no que se refere a explicar o valor de mercado das empresas ao redor do globo. Nesse sentido, o índice de liberdade econômica, mensurado pela The Heritage Foundation (2024), concebido para refletir o ambiente econômico e empresarial em cada país estudado da forma mais equilibrada possível, abrange fatores quantitativos e qualitativos, agrupados em quatro grandes pilares, de liberdade econômica: (i) Estado de Direito; (ii) Tamanho do Governo; (iii) Eficiência Regulatória; e (iv) Abertura dos Mercados. Estudos têm indicado relação entre esse índice e o desempenho (e conseqüentemente, o valor corporativo) de empresas ao redor do mundo (e.g., Chortareas et al., 2013; Lin et al., 2016; Asteriou et al., 2021). Assim, foi proposta a Hipótese 4:

H₄ – O maior nível de liberdade econômica afeta positiva e significativamente o desempenho dos modelos baseados em algoritmos de IA.

Como as informações provenientes das demonstrações contábeis divulgadas periodicamente pelas organizações são relevantes informações públicas precificadas pelo mercado, conforme Al Ani (2021), analistas, credores, gestores e outros agentes econômicos almejam averiguar a qualidade dessas informações em mercados desenvolvidos e emergentes para o desenvolvimento de modelos preditivos. Nesse contexto, Chaney et al. (2011) e Boina e Macedo (2018) destacaram que a qualidade da informação contábil divulgada pelas empresas pode ser mensurada de diversas formas. Al-Enzy et al. (2023) ressaltam que a convergência às normas internacionais de Contabilidade (*International Financial Reporting Standards – IFRS*) tem o potencial de melhorar a qualidade da informação contábil, o que tem um efeito na precificação dos ativos. Dessa forma, foi proposta a Hipótese 5:

H₅ – A adoção das IFRS afeta positiva e significativamente o desempenho dos modelos baseados em algoritmos de IA.

Por fim, destaca-se que o valor das empresas no mercado pode ser bastante influenciado por crises. Recentemente, o mundo foi afetado por duas crises: Pandemia de Covid-19 e a guerra entre Rússia e Ucrânia. No que se refere à pandemia, cuja rápida expansão mundial ocorreu no primeiro semestre de 2020, Kaczmarek et al. (2021), Ichsan et al. (2021), Mulyaningsih et al. (2021) e Avelar et al. (2022) evidenciaram efeitos diversos negativos sobre o desempenho e o valor de empresas globalmente, sendo que alguns setores foram mais impactados que outros. No que tange à guerra entre Rússia e Ucrânia, que se iniciou com a invasão russa do território ucraniano em fevereiro de 2022, esta gerou uma crise que afetou de forma significativa empresas ao redor do mundo. Muitos estudos relataram resultados corporativos negativos após a eclosão desse conflito, particularmente em nações mais expostas aos países beligerantes (e.g., Bougias et al., 2022; Marcinkowska, 2022; Tosun & Eshraghi, 2022; French et al., 2023), com apenas algumas empresas em setores e países específicos permanecendo relativamente imunes a tais efeitos adversos (Abbassi et al., 2023).

As referidas crises afetaram significativamente, inclusive, as empresas do grupo de países mais industrializados do mundo, o G20. Fundado em 1999 após a crise financeira asiática, o G20 atua como um fórum para os Ministros das Finanças e Presidentes dos Bancos

Centrais discutirem questões econômico-financeiras globais. É formado pela União Europeia, União Africana e por 19 países, quais sejam: Argentina, Austrália, Brasil, Canadá, China, França, Alemanha, Índia, Indonésia, Itália, Japão, República da Coreia, México, Rússia, Arábia Saudita, África do Sul, Turquia, Reino Unido e Estados Unidos da América (EUA). Esses países representam aproximadamente 85% do produto interno bruto (PIB) global, mais de 75% do comércio mundial e cerca de dois terços da população (G20, 2024). Assim, foi proposta a Hipótese 6:

H₆ – Crises afetam negativa e significativamente o desempenho dos modelos baseados em algoritmos de IA.

Na Figura 1, ilustram-se as principais variáveis e hipóteses que foram testadas na pesquisa ora apresentada, conforme os procedimentos metodológicos descritos na seção a seguir.

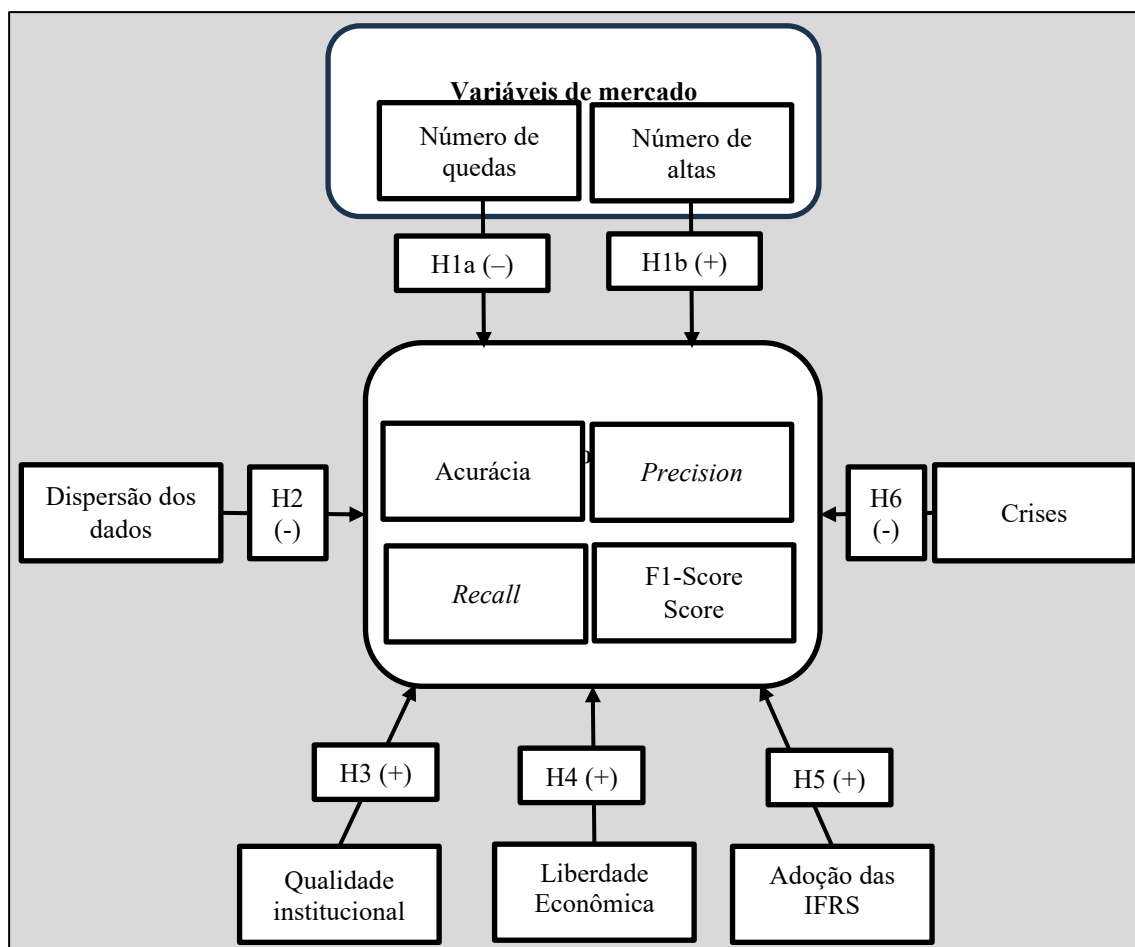


Figura 1. Principais variáveis e hipóteses testadas.

Fonte: Elaboração própria dos autores.

3. Método de pesquisa

O estudo apresentado neste artigo teve caráter descritivo e quantitativo, com base na classificação de Cooper e Schindler (2006). Nesse sentido, foram analisados dados econômico-financeiros coletados em mercados de capitais dos países que compõe o G20 mencionados anteriormente. Destaca-se que a União Europeia e a União Africana, embora façam parte do grupo, não foram consideradas na pesquisa, pois não são países *stricto sensu*. A pesquisa foi desenvolvida em duas etapas: (1) desenvolvimento dos modelos baseados nos algoritmos de IA; e (2) identificação dos determinantes do desempenho desses modelos.

3.1 Desenvolvimento dos modelos de IA

Inicialmente, foram identificados os principais índices das bolsas de valores de cada país e coletado seu *ticker* (código) no site Yahoo!Finance: “^MERV”, “^AXJO”, “^BVSP”, “^GSPTSE”, “000001.SS”, “^FCHI”, “^GDAXI”, “^BSESN”, “^JKSE”, “FTSEMIB.MI”, “^N225”, “^KS11”, “^MXX”, “IMOEX.ME”, “^TASI.SR”, “^XU100”, “^FTSE” e “^GSPC”. A informação de *tickers* foi essencial para coleta de dados referentes às cotações diárias de cada índice a partir da linguagem de programação R e das funções do pacote *Quantitative Financial Modelling Framework* (quantmod). As cotações desses índices foram coletadas em frequência diária entre os anos de 2001 e 2024. Tais dados foram tratados e usados para calcular os indicadores técnicos. Salienta-se que esses indicadores foram normalizados para possibilitar o emprego de algoritmos que demandam variáveis contínuas, tal como recomendado e implementado em estudos similares (e.g., Hsu et al., 2016; Ampomah et al., 2021; Mohanty et al., 2021). Os indicadores técnicos foram empregados para prever o retorno do índice no dia seguinte, ou seja, aumento ou redução do seu preço em relação ao dia anterior, segundo modelo apresentado na Equação 5, onde RA representa o retorno do ativo em relação ao dia anterior (1 se positivo, 0 se negativo), t representa o dia e ε representa o erro.

$$RA_t = \beta_0 + \beta_1 SMA_{t-1} + \beta_2 WMA_{t-1} + \beta_3 MACD_{t-1} + \beta_4 STCK_{t-1} + \beta_5 STCD_{t-1} + \beta_6 RSI_{t-1} + \beta_7 WPR_{t-1} + \beta_8 CCI_{t-1} + \varepsilon \quad (5)$$

Inicialmente, foram estimados modelos de Regressão Logística (RL) para servir de *benchmarking* para os algoritmos de IA, em harmonia com Chernavin (2019) e Farahani et al. (2023). Todos os modelos gerados por algoritmos de IA analisados na pesquisa (KNN, NB e RF) foram empregados para fins de classificação, com o propósito de prever o retorno futuro dos índices, usando os mesmos indicadores técnicos listados na seção anterior, em consonância com os estudos de Hsu et al. (2016), Chernavin (2019), Aminimehr et al. (2020), Ampomah et al. (2021), Farahani et al. (2023) e Srivastava et al. (2023). Os modelos passaram por etapas de treinamento e teste, sendo empregados 80% dos dados na primeira etapa e os 20% remanescentes na segunda, em linha com proporções utilizadas em estudos anteriores para previsão de preços no mercado financeiro (e.g., Mndawe et al., 2022; Srivastava et al., 2023; Shaban et al., 2024).

Para otimizar os parâmetros dos modelos e estimar o desempenho de cada um, optou-se pelas principais métricas de aprendizado de máquina para fins de classificação previamente citadas (equações de 1 a 4) em consonância com estudos pretéritos, tais como Chernavin (2019), Mohanty et al. (2021), Mndawe et al. (2022), Farahani et al. (2023) e Sagaceta-Mejía et al. (2024). Destaca-se que alguns parâmetros específicos foram ajustados para determinados

algoritmos, possibilitando sua otimização dos modelos, conforme Hsu et al. (2016), Khoa e Huynh (2021), Farahani et al. (2023) e Avelar e Jordão (2024): (a) KNN – possibilitou-se uma variação entre 1 (um) e 50 (cinquenta) para o parâmetro “k”, de modo que o algoritmo selecionasse o melhor “k-vizinhos mais próximos” para cada índice; e (b) RF – possibilitou-se uma variação entre 3 (três) e 50 (cinquenta) árvores estimadas, para que o algoritmo selecionasse a quantidade que forneceria o melhor desempenho para cada índice. Acentua-se que todos os indicadores técnicos foram calculados para o período entre 3 (três) e 30 (dias). Tal medida possibilitou que cada algoritmo selecionasse a janela com melhor desempenho para previsão de cada índice, otimizando seu desempenho, tendo em vista o exposto em Hsu et al. (2016), Khoa e Huynh (2021), Srivastava et al. (2023) e Gupta e Kumar (2024).

Ressalta-se a forma de treinamento e teste dos modelos, foi realizada com base nas etapas de Ferreira et al. (2021) para previsão de preços de ativos financeiros usando algoritmos de IA: (1) Coleta dos dados de entrada – dados dos preços históricos (abertura, máximo, mínimo e fechamento) dos principais índices de bolsas de valores selecionadas; (2) Transformação e seleção dos dados – cálculo dos retornos diários dos índices e dos respectivos indicadores técnicos; (3) Treinamento do modelo – uso de 80% dos dados por ano; (4) Otimização de parâmetros – variação da janela de cálculo dos indicadores técnicos, assim como de parâmetros específicos de cada algoritmo; e (5) Avaliação do desempenho do modelo – emprego das métricas citadas nas equações de 1 a 4.

Nessa etapa, todos os dados foram tratados e analisados por meio do *software* R, utilizando os seguintes pacotes: *Discrete Goodness-of-Fit Tests* (dgof); *quantmod*; *Functions for Classification* (class); *Breiman and Cutler's Random Forests for Classification and Regression* (randomForest); *Misc Functions of the Department of Statistics, Probability Theory Group* (e1071); *A Grammar of Data Manipulation* (dplyr); *eXtensible Time Series* (xts); e *Training of Neural Networks* (neuralnet).

3.2 Identificação dos determinantes do desempenho

Após o desenvolvimento dos modelos baseados em algoritmos de IA e da mensuração de seu desempenho, foram usadas regressões com dados em painel para analisar os determinantes das performances obtidas por tais modelos em relação a variáveis determinantes. No total, foram 16 (dezesesseis) variáveis dependentes, correspondentes às quatro métricas de avaliação de desempenho (acurácia, precisão, *Recall* e F1-Score) para cada algoritmo (RL, NB, RF e KNN). Para explicação desse desempenho, foram inicialmente definidas 32 variáveis para a estimação dos modelos determinantes do desempenho do de algoritmos de IA segregados em cinco grandes dimensões, conforme a Tabela 1.

Tabela 1: Variáveis independentes por dimensão

| Dimensão | Variáveis testadas | Referências |
|----------------------------|--|--|
| Características do mercado | Produto interno bruto (PIB), Retorno do mercado, Número de quedas e Número de aumentos | Shynkevich et al. (2017), Wang et al. (2018), Wu et al. (2020), Qian e Rasheed (2007), Patel et al. (2015), Ampomah et al. (2021), Awan et al. (2021) e Abbassi et al. (2023). |
| Variáveis institucionais | <i>International Financial Reporting Standards</i> (IFRS); Sistema legal; controle de corrupção (CC); Eficácia do governo (EG); Estabilidade | The World Bank (2024), Al-Enzy et al. (2023), Kaufmann et al. (2011), Kaufmann e Kraay (2023), |

| | | |
|-------------------------------|--|--|
| | política e ausência de violência/terrorismo (EPAV); Qualidade regulatória (QR); Estado de direito (ED); Voz e responsabilidade (VR). | Saona e Martín (2018), Ojeka et al. (2019), Thakur et al. (2021) e Bagh et al. (2023). |
| Características dos dados | Média, desvio-padrão, coeficiente de variação e número de observações | Anysz et al. (2016), Novak e Velušček (2016), Lantz (2019), Burguer (2018) e Nwanganga e Chapple (2020) |
| Índice de liberdade econômica | Pontuação geral (PG); direitos de propriedade (DP); integridade do governo (IG); eficácia judicial (EJ); gastos do governo (GG); carga tributária (CT); saúde fiscal (SF); liberdade de negócios (LN); liberdade trabalhista (LT); liberdade monetária (LM); liberdade de comércio (LC); liberdade de investimento (LI); liberdade financeira (LF) | The Heritage Foundation (2024), Chortareas et al. (2013), Lin et al. (2016) e Asteriou et al. (2021). |
| Crises | Crise do subprime, pandemia da COVID-19 e guerra entre Rússia e Ucrânia | Kaczmarek et al. (2021), Ichsan et al. (2021), Mulyaningsih et al. (2021), Avelar et al. (2022), Bougias et al. (2022), Marcinkowska (2022), Tosun e Eshraghi, (2022), Abbassi et al. (2023) e French et al. (2023). |

Fonte: Elaborado pelos autores, 2025.

A partir de essas variáveis, foram selecionadas aquelas que apresentaram maior poder explicativo em análises exploratórias, quais sejam: número de quedas, número de altas, coeficiente de variação, indicador ED de qualidade institucional (QI_ED_d), indicador PG de liberdade econômica (IEF_PG_d) e crises macroeconômicas (d_crises). A Tabela 2 resume detalhes das variáveis independentes. Assim, a especificação formal do modelo geral está apresentada na Equação 6.

Tabela 2: Definição das variáveis independentes do Modelo de Análise de Regressão com Dados em Painel

| Sigla | Definição | Hipótese |
|--------------|--|-----------------|
| Num quedas | Número de quedas (retornos negativos) diárias do índice | H _{1a} |
| Num aumentos | Número de aumentos (retornos positivos) diários do índice | H _{1b} |
| CV | Coeficiente de variação | H ₂ |
| QI_ED_d | Estado de Direito – resume das percepções sobre até que ponto os agentes confiam e cumprem as regras da sociedade e, em particular, a qualidade da execução de contratos, os direitos de propriedade, a polícia e os tribunais, bem como a probabilidade de crime e violência. | H ₃ |
| IEF_PG_d | Pontuação geral do Índice de Liberdade Econômica – Resume todos os grupos desse índice. | H ₄ |
| d_IFRS | <i>Dummy</i> que identifica a partir de que ano o país aderiu ao IFRS, do inglês <i>International Financial Reporting Standards</i> em seu sistema contábil nacional. | H ₅ |
| d_crise | <i>Dummy</i> que identifica o ano da crise de 2008 (subprime), de 2020 (COVID-19) e 2022 (guerra Rússia e Ucrânia). | H ₆ |

Fonte: Elaborado pelos autores, 2025.

$$Medida_ML_{it}^n = \beta_0 + \beta_1 d_IFRS_{it} + \beta_2 Num_quedas_{it} + \beta_3 Num_aumentos_{it} + \beta_4 CV_{it} + \beta_5 QI_ED_d_{it} + \beta_6 IEF_PG_d_{it} + \beta_7 d_crise_t + \varepsilon_{it} \quad (6)$$

Onde: $Medida_ML_{it}^n$ refere-se a cada uma das n medidas de desempenho do modelo de IA, com $n = 1, \dots, 16$, são elas: $Acuracia_NB$, $Precision_NB$, $Recall_NB$, $F1-Score_NB$, $Acuracia_RL$, $Precision_RL$, $Recall_RL$, $F1-Score_RL$, $Acuracia_RF$, $Precision_RF$, $Recall_RF$, $F1-Score_RF$, $Acuracia_KNN$, $Precision_KNN$, $Recall_KNN$ e $F1-Score_KNN$. β_0 é o intercepto; os β_1, \dots, β_7 são os coeficientes estimados; os subscritos i e t referem-se a país e ano, respectivamente; e ε é o termo de erro aleatório.

Em todas as 16 estimações, adotou-se um modelo de dados em painel com efeitos fixos. Essa escolha se justifica pelo pressuposto de que a heterogeneidade não observada de cada país analisado é constante ao longo do tempo, mas ainda assim pode influenciar os resultados das estimações (Fávero, 2013). Desse modo, características específicas de cada país, que não mudam significativamente ao longo do período de análise e usualmente não são capturadas pelas variáveis incluídas no modelo (como cultura, instituições políticas históricas, etc.), podem influenciar as variáveis dependentes e independentes. Se essa heterogeneidade não observada for correlacionada com as variáveis explicativas, sua omissão poderia acarretar vieses nas estimações. O modelo de efeitos fixos, ao lidar com essa heterogeneidade, possibilita maior confiabilidade nos resultados das estimações.

4. Resultados

A estatística descritiva das variáveis é apresentada na Tabela 3. As médias das variáveis de Acurácia, Precisão, *Recall* e F1-Score indicaram um alto desempenho dos modelos de IA, todas superiores a 60%. Esse resultado indica que os desempenhos dos modelos de IA são superiores ao esperado pelo mercado com base na HME na forma fraca (que seria 50%). É importante destacar que, independentemente da medida usada para avaliar o desempenho, os modelos que empregaram o algoritmo RF obtiveram o melhor desempenho na média. É importante ressaltar que a diferença no número de observações de 396 para 384 ocorreu porque a Rússia está com dados ausentes de 2002 a 2012 e Turquia em 2021. Portanto, trata-se de uma estrutura de dados desbalanceada.

Tabela 3: Estatística Descritiva das variáveis métricas

| Variável | Média | Desvio-Padrão | Mínimo | Máximo | N |
|---------------|-----------|---------------|-----------|-----------|-----|
| Acuracia NB | 0,6520158 | 0,0901486 | 0,45 | 0,9375 | 384 |
| Acuracia RL | 0,6636671 | 0,0801299 | 0,4545455 | 1 | 384 |
| Acuracia RF | 0,7839613 | 0,0573492 | 0,6176471 | 1 | 384 |
| Acuracia KNN | 0,6203897 | 0,0213853 | 0,5777778 | 0,7112676 | 384 |
| Precision NB | 0,7048445 | 0,141466 | 0,3333333 | 1 | 384 |
| Precision RL | 0,7114514 | 0,1383471 | 0,3333333 | 1 | 384 |
| Precision RF | 0,8773084 | 0,1142082 | 0,5238095 | 1 | 384 |
| Precision KNN | 0,8387792 | 0,1225717 | 0,4444444 | 1 | 384 |
| Recall NB | 0,7557972 | 0,1801753 | 0,1666667 | 1 | 384 |
| Recall RL | 0,7958981 | 0,1813351 | 0,2173913 | 1 | 384 |
| Recall RF | 0,9273535 | 0,0748817 | 0,65 | 1 | 384 |
| Recall KNN | 0,9267555 | 0,0933448 | 0,6111111 | 1 | 384 |
| F1-Score NB | 0,6781232 | 0,1149299 | 0,2666667 | 0,96 | 384 |

| Variável | Média | Desvio-Padrão | Mínimo | Máximo | N |
|--------------|-----------|---------------|-----------|-----------|-----|
| F1-Score RL | 0,6938632 | 0,1077189 | 0,3333333 | 1 | 384 |
| F1-Score RF | 0,8048668 | 0,0650516 | 0,625 | 1 | 384 |
| F1-Score KNN | 0,7941814 | 0,0725342 | 0,5454545 | 1 | 384 |
| Num quedas | 115,2109 | 11,20611 | 69 | 143 | 384 |
| Num aumentos | 130,1927 | 12,64826 | 75 | 170 | 384 |
| CV | 0,0865648 | 0,0727603 | 0,0155774 | 0,8786258 | 384 |
| IEF PG d | 0,6484545 | 0,0974245 | 0,438 | 0,831 | 396 |
| QI ED d | 0,6017808 | 0,2861124 | 0,0190476 | 0,9668247 | 396 |

Nota: N refere-se ao número de observações.

Fonte: Elaborado pelos autores, 2025.

Os resultados das regressões em painel com efeito fixos para os determinantes de IA estão apresentados na Tabela 4. Os resultados para os determinantes da RL demonstraram que: (i) número de quedas apresentou efeito negativo e estatisticamente significativo em todos os modelos, exceto em Precisão; (ii) o número de aumentos apresentou efeito positivo e significativo nos modelos de *Recall* e F1-Score, mas exerceu efeito negativo no modelo de Acurácia; (iii) as variáveis “CV” e “Crise” foram significativas estatisticamente no modelo *Recall*, sendo na primeira de forma negativa e a segunda de forma positiva.

Tabela 4: Resultado dos modelos de determinantes de desempenho da IA

| Painel A | | | | | | | | |
|----------------|-------------|-----------|-------------|-------------|-----------|-----------|-------------|-------------|
| Algoritmo | RL | | | | NB | | | |
| | Acurácia | Precision | Recall | F1-Score | Acurácia | Precision | Recall | F1-Score |
| d_IFRS | -0,0258** | -0,0138 | -0,0314 | -0,0211* | -0,0111 | 0,00724 | -0,0491* | -0,0163 |
| Quedas | -0,00324*** | -0,00175 | -0,00630*** | -0,00506*** | -0,00106* | 0,000696 | -0,00542*** | -0,00308*** |
| Aumentos | -0,00181* | -0,00226 | 0,00580*** | 0,00194** | 0,000180 | -0,000265 | 0,00581*** | 0,00379*** |
| CV | -0,0387 | 0,0826 | -0,179** | -0,0968 | -0,0653 | 0,0816 | -0,205** | -0,131 |
| IEF_PG_d | -0,0903 | 0,0557 | -0,147 | -0,0541 | 0,178 | 0,0712 | 0,194 | 0,140 |
| QI_ED_d | -0,00273 | 0,0112 | 0,0827 | -0,00992 | -0,175** | -0,173 | -0,00119 | -0,187* |
| d_crise | 0,0148 | -0,0206 | 0,0294* | 0,00297 | 0,00494 | -0,0370* | -0,00363 | -0,00920 |
| Constante | 1,348*** | 1,167** | 0,840** | 1,085*** | 0,753*** | 0,713* | 0,541** | 0,584** |
| R ² | 0,069 | 0,015 | 0,487 | 0,460 | 0,032 | 0,012 | 0,451 | 0,418 |

| Painel B | | | | | | | | |
|----------------|------------|-----------|-------------|-------------|------------|------------|-------------|-------------|
| Algoritmo | RF | | | | KNN | | | |
| | Acurácia | Precision | Recall | F1-Score | Acurácia | Precision | Recall | F1-Score |
| d_IFRS | 0,00372 | 0,00892 | -0,00485 | 0,00229 | -0,00162 | -0,00353 | -0,00762 | 0,00367 |
| Quedas | -0,00124** | -0,000980 | -0,00262*** | -0,00219*** | -0,00059** | 0,000454 | -0,00405*** | -0,00270*** |
| Aumentos | -0,00110** | -0,00204* | -2,18e-05 | -5,25e-05 | 6,00e-05 | -0,00225** | 0,00136** | 0,000833** |
| CV | 0,00453 | 0,0187 | 0,00459 | 0,0124 | 0,0620** | 0,0580 | -0,0532 | -0,0718 |
| IEF_PG_d | -0,0667 | 0,294** | -0,179 | -0,0171 | -0,0283 | 0,0674 | 0,0255 | 0,0375 |
| QI_ED_d | -0,0590 | -0,253** | 0,0805 | -0,0338 | 0,0165 | -0,0286 | 0,0963* | -0,0299 |
| d_crise | 0,0131 | -0,0249 | 0,0116 | 0,000346 | 0,00146 | -0,00590 | 0,0183 | 0,00166 |
| Constante | 1,146*** | 1,216*** | 1,300*** | 1,094*** | 0,684*** | 1,051*** | 1,148*** | 0,995*** |
| R ² | 0,023 | 0,035 | 0,141 | 0,128 | 0,166 | 0,062 | 0,359 | 0,264 |

Fonte: Elaborado pelos autores, 2025.

Níveis de significância: * 10 %, ** 5% e *** 1%.

No caso dos modelos baseados no algoritmo NB, o número de quedas foi o determinante mais consistente nas estimações, apresentando influência negativa e significativa estatisticamente em todos os modelos, exceto no modelo Precisão. Por outro lado, o número de aumentos exerceu influência positiva e significativa, mas somente nos modelos de *Recall* e F1-Score. A qualidade institucional, por sua vez, exerceu influência negativa e significativa nos modelos de Acurácia e F1-Score. Adicionalmente, as variáveis que influenciaram negativa e significativa estatisticamente em apenas um dos modelos foram: IFRS, em *Recall*; CV, em *Recall*; e crise, em Precisão.

Novamente, os resultados encontrados para as estimações dos modelos NB e RL foram similares comparativamente com o algoritmo RF para algumas variáveis. Nesse caso, os principais destaques foram: (i) número de quedas influenciou de forma negativa e com significância estatística em todos os modelos, exceto no modelo de Precisão; (ii) o número de aumentos apresentou uma relação negativa nos modelos Acurácia e Precisão; (iii) o índice de liberdade econômica influenciou positivamente e de forma significativa estatisticamente, no modelo de Precisão; e (iv) a qualidade institucional influenciou negativa e significativamente, no modelo de Precisão. Diferentemente do que foi encontrado nos modelos RL e NB, as variáveis *dummy* de IFRS e crise não foram significativas estatisticamente.

Por fim, os resultados encontrados para as estimações com o algoritmo KNN foram similares aos apresentados anteriormente: (a) número de quedas influenciou de forma inversamente proporcional e com significância estatística todos os modelos, exceto o de Precisão; (b) o número de aumentos foi diretamente proporcional nos modelos *Recall* e F1-Score, a diferença é que neste caso o coeficiente foi inversamente proporcional no modelo de Precisão; (c) as variáveis CV e qualidade institucional influenciaram de forma positiva e estatisticamente significativa nos modelos de Acurácia e *Recall*, respectivamente.

Diante do exposto, verifica-se que, de um lado, as variáveis determinantes dos modelos de IA não foram significativas no conjunto dos 16 modelos estimados. Em contrapartida, é possível perceber que os determinantes exerceram alguma influência de forma isolada e alternativa em cada uma das quatro métricas de avaliação de desempenho — acurácia, precisão, *Recall* e F1-Score—, calculada a partir de quatro algoritmos de IA diferentes — RL, NB, RF e KNN.

4. Discussão

Com base na primeira etapa da pesquisa, observou-se que o desempenho dos modelos baseados em algoritmos de IA foi superior (60%) ao desempenho esperado com base na HME, notadamente em sua forma fraca, em consonância com estudos desenvolvidos na última década, como Shynkevich et al. (2017), Cao et al. (2019), Chernavin (2019), Mndawe et al. (2022) e Sagaceta-Mejía et al. (2024). Salienta-se que os modelos que apresentaram as maiores médias de desempenho (independentemente da métrica usada para mensurá-lo) sempre foram aqueles baseados no algoritmo RF. Esse resultado sinaliza a robustez e eficácia do algoritmo RF em capturar padrões nos dados que outros algoritmos não conseguem. Também válida o uso de algoritmos de IA para superar *benchmarks* estabelecidos, como a HME, o que é desejável para aplicações práticas na qual a precisão é fundamental para os tomadores de decisões.

Observou-se que todos os modelos estimados foram estatisticamente significantes, sendo que houve diferenças nas variáveis determinantes de acordo com a métrica de desempenho empregada. Tal constatação evidencia que as variáveis podem explicar as

diferenças de desempenho observadas quando diferentes métricas são empregadas para mensurá-lo, tal como reportado por Patel et al. (2015), Nabipour et al. (2017), Wang et al. (2018), Wu et al. (2020) e Avelar e Jordão (2024). Nesse sentido, os níveis de explicação dos modelos *Recall* e F1-Score foram sempre superiores aos níveis daqueles dos modelos Acurácia e Precisão. Assim, há evidências de que, independentemente do algoritmo de IA empregado, as variáveis determinantes propostas neste estudo tendem a explicar o desempenho dos modelos conforme as métricas empregadas.

Sobre os determinantes do desempenho, especificamente, a principal variável que apresentou resultados consistentemente negativos e significantes em todos os modelos (com exceção daqueles cuja variável de desempenho era a precisão) foi aquela relacionada ao número de retornos negativos (quedas) no período. Trata-se de uma variável essencial para o cálculo de desempenho desses modelos, tendo em vista Huang et al. (2008), Nabipour et al. (2017), Moon et al. (2018), Cao et al. (2019), Awan et al. (2021) e Hung et al. (2024), sendo enfatizada de diferentes formas nas principais métricas de aprendizado de máquina usadas para fenômenos de classificação (Burguer, 2018; Lantz, 2019; Nwanganga & Chapple, 2020). Dessa forma, há evidências robustas em favor da H1a. Saliencia-se que a variável de aumento (retorno positivo) também foi significativa na maior parte dos modelos, porém, sem uma consistência como a variável anterior. Destaca-se, porém, que, apesar de essa variável ser positiva na maior parte dos modelos em que foi significativa, ela foi significativa e negativa em dois modelos que empregaram a métrica Precisão como forma de mensurar o desempenho. Assim, pode-se dizer que há evidências limitadas (menos consistentes) em favor da H1b.

Por sua vez, as demais variáveis (IEF_PG_d, QI_ED_d, CV e d_crise) se mostraram significantes em alguns modelos, porém, sem consistência nos sinais dos coeficientes e nas métricas empregadas para avaliar o desempenho. Embora a variável IFRS tenha sido negativa nos modelos em que foi significativa, houve poucas situações como essa. Assim, os resultados obtidos não seguiram um padrão, podendo ser considerados contextuais, não gerando evidências robustas para explicação do fenômeno. Dessa maneira, não foi possível ratificar as demais hipóteses do estudo, indo de encontro ao esperado na literatura sobre modelos de aprendizado de máquina (e.g., Anysz et al., 2016; Novak & Velušček, 2016; Lantz, 2019; Nwanganga & Chapple, 2020; Ali, 2022; Imani et al., 2024) e na literatura sobre preditores de valor corporativo (e.g., Chortareas et al., 2013; Lin et al., 2016; Saona & Martín, 2018; Ojeka et al. 2019; Thakur et al., 2021; Bagh et al., 2023; Asteriou et al., 2021; Ichsan et al., 2021; Kaczmarek et al., 2021; Mulyaningsih et al.; 2021; Avelar et al.; 2022; Bougias et al., 2022; Marcinkowska, 2022; Tosun & Eshraghi, 2022; Abbassi et al., 2023; Al-Enzy et al., 2023; French et al., 2023).

Tais resultados trazem contribuições importantes sob diversas perspectivas. Sob a perspectiva teórica, o estudo traz novas evidências robustas sobre como o emprego de modelos baseados em algoritmos de IA podem, consistente e sistematicamente, obter um desempenho superior ao esperado com base no mercado de acordo com a HME (em especial, na sua forma fraca). Ademais, dentre tais algoritmos, o estudo destaca o desempenho constantemente superior dos modelos baseados em RF em relação aos demais. Por outro lado, esse achado reforça a discussão sobre um possível *trade-off* entre transparência e desempenho dos modelos de IA para previsão de fenômeno em finanças, tal como já evidenciado em Avelar e Jordão (2024). Além disso, o estudo realça o poder preditivo oferecido pela IA aos indicadores técnicos para previsão de preços no mercado financeiro.

Sob a perspectiva da literatura em finanças corporativas, trata-se de um estudo inovador, pois é o primeiro a focar as variáveis que afetam o desempenho dos modelos baseados em algoritmos de IA, apesar de seu amplo uso para previsão na área nas últimas décadas. Em especial, foram identificadas diversas variáveis que potencialmente podem afetar o desempenho desses modelos e foram obtidas evidências robustas de que o desempenho dos modelos preditivos diminui quando há um número maior de retornos negativos (quedas). No que se refere às outras variáveis, constatou-se um comportamento contextual, indicando que as variáveis tradicionalmente destacadas em estudos sob valor corporativo podem ser menos consistentes para explicar tais variações no âmbito do uso de modelos baseados em IA.

Diante do exposto, considerando a perspectiva dos agentes de mercado (investidores, analistas, financiadores etc.), os resultados reforçaram a necessidade de analisar o contexto, uma vez que um mercado com maior número de quedas dos preços de ativos financeiros tende a piorar o desempenho de modelos preditivos que usam algoritmos de IA. Ademais, outras variáveis, apesar de resultados menos consistentes, tenderam a afetar determinadas métricas de desempenho dependendo do algoritmo de IA empregado. Por fim, destacou-se ainda a superioridade do desempenho do algoritmo RF sobre os demais, sendo uma opção mais promissora (dentre as analisadas) para agentes de mercado.

5. Conclusões

Este artigo apresentou os resultados de um estudo que visou analisar os determinantes do desempenho dos modelos baseados em algoritmos de IA para a previsão de preços dos principais índices das bolsas de valores de países membros do G20 ao longo do século XXI. Constatou-se que, não só o desempenho dos modelos baseados em algoritmos de IA foi superior ao desempenho esperado pelo mercado, variando de acordo com a métrica empregada, como também que o desempenho dos modelos preditivos diminuiu quando há um número maior de retornos negativos (quedas) no período analisado. Os resultados também realçaram o poder preditivo oferecido pela IA aos indicadores técnicos para previsão de preços no mercado financeiro. Destacou-se, ainda, a superioridade do desempenho do algoritmo RF sobre os demais, evidenciando um potencial *trade-off* entre desempenho e transparência dos algoritmos.

Salienta-se que o estudo apresentou uma série de contribuições sob diferentes perspectivas: (i) ineditismo em analisar variáveis determinantes do desempenho de modelos baseados em algoritmos de IA; (ii) evidências robustas sobre desempenho consistente e sistematicamente superior desses modelos baseados em relação ao mercado; e (iii) poder preditivo oferecido por algoritmos de IA aos indicadores técnicos para previsão de preços.

Ainda assim, algumas limitações foram encontradas. Dentre elas, destacaram-se: (a) outros determinantes podem influenciar nas relações estimadas, mas que não foram incluídas nos modelos; (b) outros algoritmos de IA também poderiam ter sido utilizados; e (c) países em diferentes estágios de desenvolvimento podem ter contribuído para a heterogeneidade dos resultados encontrados. Como sugestão para estudos futuros, recomenda-se: (i) a divisão de países de acordo com características semelhantes; (ii) proposição de mais variáveis determinantes, tais como informações mais específicas e desagregadas dos países; (iii) inclusão de outros tipos de algoritmos de IA e de medições; e (iv) desenvolvimento de outras formas de análise dos modelos baseados em algoritmos de IA e seus determinantes.

Referências

- Abbassi, W., Kumari, V., & Pandey, D. K. (2023). What makes firms vulnerable to the Russia–Ukraine crisis? *Journal of Risk Finance*, 24(1), 24–39. <https://doi.org/10.1108/JRF-05-2022-0108>
- Al Ani, M. K. (2021). Earnings quality and foreign investors in gulf cooperation council countries. *Contemporary Management Research*, 17(3), 223–270.
- Al-Enzy, N. S. K., Monem, R., & Nahar, S. (2023). IFRS experience and earnings quality in the GCC region. *International Journal of Managerial Finance*, 9(3), 670–690. <https://doi.org/10.1108/IJMF-09-2021-0410>
- Ali, P. J. M. (2022). Investigating the Impact of min-max data normalization on the regression performance of K-nearest neighbor with different similarity measurements. *ARO-The Scientific Journal of Koya University*, 10(1), 85-91. <https://aro.koyauniversity.org/index.php/aro/article/view/955/262>
- Ampomah, E. K., Qin, Z., Nyame, G., & Botchey, F. E. (2021). Stock market decision support modeling with tree-based AdaBoost ensemble machine learning models. *Informatica*, 44(4). <https://doi.org/10.31449/inf.v44i4.3159>
- Anysz, H., Zbiciak, A., & Ibadov, N. (2016). The influence of input data standardization method on prediction accuracy of artificial neural networks. *Procedia Engineering*, 153, 66-70. <https://doi.org/10.1016/j.proeng.2016.08.081>
- Asteriou, D., Pilbeam, K., & Tomuleasa, I. (2021). The impact of corruption, economic freedom, regulation and transparency on bank profitability and bank stability: Evidence from the Eurozone area. *Journal of Economic Behavior & Organization*, 184, 150–177. <https://doi.org/10.1016/j.jebo.2020.08.023>
- Avelar, E. A., & Jordão, R. V. D. (2024). The role of artificial intelligence in the decision-making process: A study on the financial analysis and movement forecasting of the world's largest stock exchanges. *Management Decision*. <https://doi.org/10.1108/MD-09-2023-1625>
- Avelar, E. A., Orefici, J. B. P., Borges, S. L., Campos, O. V., & Ferreira, P. O. (2022). The Covid-19 pandemic effects on the financial performance of Brazilian listed companies. *Revista de Administração da UFES*, 15(3), 389–410. <https://doi.org/10.5902/1983465967372>
- Awan, M. J., Rahim, M. S. M., Nobanee, H., Munawar, A., Yasin, A., & Zain, A. M. (2021). Social media and stock market prediction: A big data approach. *Computers, Materials & Continua*, 67(2), 2569–2583. <https://doi.org/10.32604/cmc.2021.014253>
- Bagh, T., Naseer, M. M., Khan, M. A., Pypłacz, P., & Oláh, J. (2023). Sustainable growth rate, corporate value of US firms within capital and labor market distortions: The moderating effect

of institutional quality. *Oeconomia Copernicana*, 14(4), 1211–1255.
<https://doi.org/10.24136/oc.2023.036>

Beniwal, M., Singh, A., & Kumar, N. (2024). Forecasting multistep daily stock prices for long-term investment decisions: A study of deep learning models on global indices. *Engineering Applications of Artificial Intelligence*, 129, 107617.
<https://doi.org/10.1016/j.engappai.2023.107617>

Boina, T. M., & Macedo, M. A. da S. (2018). Capacidade preditiva de accruals antes e após as IFRS no mercado acionário brasileiro. *Revista Contabilidade & Finanças*, 29(78), 375–389.

Bougias, A., Episcopos, A., & Leledakis, G. N. (2022). Valuation of European firms during the Russia–Ukraine war. *Economics Letters*, 218, 110750.
<https://doi.org/10.1016/j.econlet.2022.110750>

Burger, S. V. (2018). *Introduction to machine learning with R: Rigorous mathematical analysis*. O'Reilly Media.

Cao, H., Lin, T., Li, Y., & Zhang, H. (2019). Stock price pattern prediction based on complex network and machine learning. *Complexity*, 2019, 4132485.
<https://doi.org/10.1155/2019/4132485>

Chaney, P. K., Faccio, M., & Parsley, D. (2011). The quality of accounting information in politically connected firms. *Journal of Accounting and Economics*, 51(1–2), 58–76.

Chernavin, N. P. (2019). Application of the committee machine method to analysis of stock market technical indicators. *Бизнес-информатика*, 13(4), 73–86.

Chortareas, G. E., Girardone, C., & Ventouri, A. (2013). Financial freedom and bank efficiency: Evidence from the European Union. *Journal of Banking & Finance*, 37(4), 1223–1231.
<https://doi.org/10.1016/j.jbankfin.2012.11.015>

Cooper, D.R., & Schindler, P.S. (2006) *Business Research Methods*. 8th Edition, McGraw Hill, Tata.

Farahani, M. S., Farrokhi-Asl, H., & Rahimian, S. (2023). Hybrid metaheuristic artificial neural networks for stock price prediction considering efficient market hypothesis. *International Journal of Research in Industrial Engineering*, 12(3), 234–272.
<https://doi.org/10.22105/riej.2023.361216.1336>

Fávero, Luiz P. *Métodos Quantitativos com Stata*. Rio de Janeiro: GEN LTC, 2013.

Ferreira, F. G., Gandomi, A. H., & Cardoso, R. T. (2021). Artificial intelligence applied to stock market trading: A review. *IEEE Access*, 9, 30898–30917.
<https://doi.org/10.1109/ACCESS.2021.3058133>

French, J. J., Gurdgiev, C., & Shin, S. (2023). When doing the right thing doesn't pay: Impact of corporate decisions on Russian market participation in the wake of the Ukraine-Russia war. *Finance Research Letters*, 58, 104468. <https://doi.org/10.1016/j.frl.2023.104468>

G20. (2024). *Building a just world and a sustainable planet*. <https://www.g20.org/en>

Hung, M. C., Chen, A. P., & Yu, W. T. (2024). AI-driven intraday trading: Applying machine learning and market activity for enhanced decision support in financial markets. *IEEE Access*. Advance online publication. <https://doi.org/10.1109/ACCESS.2024.3355446>

Hsu, M. W., Lessmann, S., Sung, M. C., Ma, T., & Johnson, J. E. (2016). Bridging the divide in financial market forecasting: Machine learners vs. financial economists. *Expert Systems with Applications*, 61, 215–234. <https://doi.org/10.1016/j.eswa.2016.05.033>

Ichsan, R. N., Suparmin, S., Yusuf, M. A., Ismal, R., & Sitompul, S. (2021). Determinant of Sharia Bank's financial performance during the Covid-19 pandemic. *Budapest International Research and Critics Institute Journal (BIRCI-Journal)*, 4(1), 398–409. <https://doi.org/10.33258/birci.v4i1.1594>

Kaczmarek, Y., Perez, K., Demir, E., & Zaremba, A. (2021). How to survive a pandemic: The corporate resiliency of travel and leisure companies to the Covid-19 outbreak. *Tourism Management*, 84, 104281. <https://doi.org/10.1016/j.tourman.2020.104281>

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The worldwide governance indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law*, 3, 220–246. <https://doi.org/10.1017/S1876404511200046>

Kaufmann, D., & Kraay, A. (2023). *Worldwide governance indicators*. World Bank. <https://www.worldbank.org/en/publication/worldwide-governance-indicators>

Khoa, B. T., & Huynh, T. T. (2021). Is it possible to earn abnormal return in an inefficient market? An approach based on machine learning in stock trading. *Computational Intelligence and Neuroscience*, 2021, 2917577. <https://doi.org/10.1155/2021/2917577>

Lantz, B. (2019). *Machine learning with R: Expert techniques for predictive modeling*. Packt Publishing.

Lin, K. L., Doan, A. T., & Doong, S. C. (2016). Changes in ownership structure and bank efficiency in Asian developing countries: The role of financial freedom. *International Review of Economics & Finance*, 43, 19–34. <https://doi.org/10.1016/j.iref.2015.10.029>

Marcinkowska, M. (2022). Multinationals in Russia and Ukraine in the face of war – The stakeholders' perspective. *Comparative Economic Research*, 25(4). <https://doi.org/10.18778/1508-2008.25.27>

Mndawe, S. T., Paul, B. S., & Doorsamy, W. (2022). Development of a stock price prediction framework for intelligent media and technical analysis. *Applied Sciences*, 12(2), 719. <https://doi.org/10.3390/app12020719>

Mohanty, D. K., Parida, A. K., & Khuntia, S. S. (2021). Financial market prediction under deep learning framework using auto encoder and kernel extreme learning machine. *Applied Soft Computing*, 99, 106898. <https://doi.org/10.1016/j.asoc.2020.106898>

Moon, K., Jun, S., & Kim, H. (2018). Speed up of the majority voting ensemble method for the prediction of stock price directions. *Economic Computation & Economic Cybernetics Studies & Research*, 52(1).

Mulyaningsih, T., Cahyadin, M., & Sarmidi, T. (2021). Firms' financial distress during the Covid-19 pandemic and fiscal incentives. *Policy Brief*. Economic Research Institute for ASEAN and East Asia. <https://seads.adb.org/report/firms-financial-distress-during-covid-19-pandemic-and-fiscal-incentives>

Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data: A comparative analysis. *IEEE Access*, 8, 150199–150212. <https://doi.org/10.1109/ACCESS.2020.3015966>

Novak, M. G., & Velušček, D. (2016). Prediction of stock price movement based on daily high prices. *Quantitative Finance*, 16(5), 793–826. <https://www.tandfonline.com/doi/abs/10.1080/14697688.2015.1070960>

Nwanganga, F., & Chapple, M. (2020). *Practical machine learning in R*. John Wiley & Sons.

Ojeka, S., Adegboye, A., Adegboye, K., Umukoro, O., Dahunsi, O., & Ozordi, E. (2019). Corruption perception, institutional quality and performance of listed companies in Nigeria. *Heliyon*, 5(10), e02569. <https://doi.org/10.1016/j.heliyon.2019.e02569>

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259–268. <https://doi.org/10.1016/j.eswa.2014.07.040>

Qian, B., & Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence*, 26, 25–33. <https://doi.org/10.1007/s10489-006-0001-7>

Sagaceta-Mejía, A. R., Sánchez-Gutiérrez, M. E., & Fresán-Figueroa, J. A. (2024). An intelligent approach for predicting stock market movements in emerging markets using optimized technical indicators and neural networks. *Economics*, 18(1), 20220073. <https://doi.org/10.1515/econ-2022-0073>

Saona, P., & San Martín, P. (2018). Determinants of firm value in Latin America: An analysis of firm attributes and institutional factors. *Review of Managerial Science*, 12, 65–112. <https://doi.org/10.1007/s11846-016-0213-0>

Shaban, W. M., Ashraf, E., & Slama, A. E. (2024). SMP-DL: A novel stock market prediction approach based on deep learning for effective trend forecasting. *Neural Computing and Applications*, 36(4), 1849–1873. <https://doi.org/10.1007/s00521-023-09179-4>

Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., & Li, Y. (2017). Forecasting price movements using technical indicators: Investigating the impact of varying input window length. *Neurocomputing*, 264, 71–88. <https://doi.org/10.1016/j.neucom.2016.11.095>

Singh, J., & Khushi, M. (2021). Feature learning for stock price prediction shows a significant role of analyst rating. *Applied System Innovation*, 4(1), 17. <https://doi.org/10.3390/asi4010017>

Srivastava, S., Pant, M., & Gupta, V. (2023). Analysis and prediction of Indian stock market: A machine-learning approach. *International Journal of System Assurance Engineering and Management*, 14(4), 1567–1585. <https://doi.org/10.1007/s13198-023-01934-z>

Thakur, B. P. S., Kannadhasan, M., Charan, P., & Gupta, C. P. (2021). Corruption and firm value: Evidence from emerging market economies. *Emerging Markets Finance and Trade*, 57(4), 1182–1197. <https://doi.org/10.1080/1540496X.2019.1613643>

The Heritage Foundation. (2024). *Index of economic freedom*. <https://www.heritage.org/index/pages/about>

The World Bank. (2024). *Worldwide governance indicators (WGI)*. <https://www.worldbank.org/en/publication/worldwide-governance-indicators/documentation>

Tosun, O. K., & Eshraghi, A. (2022). Corporate decisions in times of war: Evidence from the Russia-Ukraine conflict. *Finance Research Letters*, 48, 102920. <https://doi.org/10.1016/j.frl.2022.102920>

Weng, B., Ahmed, M. A., & Megahed, F. M. (2017). Stock market one-day ahead movement prediction using disparate data sources. *Expert Systems with Applications*, 79, 153–163. <https://doi.org/10.1016/j.eswa.2017.02.041>

Wu, D., Wang, X., Su, J., Tang, B., & Wu, T. (2020). A labeling method for financial time series prediction based on trends. *Entropy*, 22(10), 1162. <https://doi.org/10.3390/e22101162>

Agradecimento:

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio na realização da pesquisa.