



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS HUMANAS SOCIAIS E AGRÁRIAS
DEPARTAMENTO DE CIÊNCIAS SOCIAIS APLICADAS
CURSO DE ADMINISTRAÇÃO**

**APRIMORANDO PREVISÕES DE RETORNOS EM ATIVOS
FINANCEIROS: uma abordagem híbrida com Arima e o modelo TimesNet**

PEDRO LUCAS FERREIRA DA SILVA

Bananeiras
Outubro 2025

PEDRO LUCAS FERREIRA DA SILVA

**APRIMORANDO PREVISÕES DE RETORNOS EM ATIVOS
FINANCEIROS: uma abordagem híbrida com Arima e o modelo TimesNet**

Trabalho de Curso apresentado como parte dos requisitos necessários à obtenção do título de Bacharel em Administração, pelo Centro de Ciências Humanas Sociais e Agrárias, da Universidade Federal da Paraíba / UFPB.

Orientador: Gustavo Correia Xavier

Bananeiras
Outubro 2025

**Catalogação na publicação
Seção de Catalogação e Classificação**

F395a Silva, Pedro Lucas Ferreira da.
Aprimorando previsões de retornos em ativos
financeiros: uma abordagem híbrida com arima e o modelo
timesnet / Pedro Lucas Ferreira da Silva. - Bananeiras,
2025.
43 f. : il.

Orientação: Gustavo Correia Xavier.
Graduação (Bacharelado em Administração) -
UFPB/CCHSA.

1. Séries Temporais. 2. ARIMA. 3. Inteligência
Artificial. 4. Deep Learning. 5. TimesNet. 6. Mercado
Financeiro. I. Xavier, Gustavo Correia. II. Título.

UFPB/CCHSA/BSMSV

CDU 658 (043)

PEDRO LUCAS FERREIRA DA SILVA

**APRIMORANDO PREVISÕES DE RETORNOS EM ATIVOS
FINANCEIROS: uma abordagem híbrida com Arima e o modelo TimesNet**

Trabalho de Curso apresentado à banca examinadora como parte do requisito parcial à obtenção do título de Bacharel em Administração, pelo Centro de Ciências Humanas Sociais e Agrárias, da Universidade Federal da Paraíba / UFPB.

Trabalho de Conclusão de Curso, Julgado e Aprovado em 13 de outubro de 2025

Dr. Gustavo Correia Xavier
Orientador



Prof. Dr. Dalton Cézane Gomes Valadares
Examinador

AGRADECIMENTOS

A trajetória até aqui foi marcada por desafios, aprendizados e transformações que transcendem os limites da sala de aula. Finalizar este trabalho simboliza mais do que o encerramento de uma etapa acadêmica, é o resultado de uma jornada construída com dedicação, resiliência e propósito.

Manifesto minha gratidão, primeiramente, à minha família, pelo apoio incondicional, pela paciência e pelo amor que sempre me sustentaram nos momentos mais difíceis. Cada palavra de incentivo e cada gesto de confiança foram essenciais para que eu acreditasse em mim mesmo e perseverasse, mesmo diante das adversidades.

Aos amigos que estiveram ao meu lado ao longo dessa caminhada, expresso meu profundo agradecimento pelas conversas, pelas risadas e pelo companheirismo que tornaram essa jornada mais leve e significativa. Em especial aos meus amigos: Kauã Herculano, Pablo Ruan, Douglas Borges e João Pedro, deixo um agradecimento especial.

Aos professores e colegas que tive a honra de conhecer ao longo do curso de Administração, expresso minha gratidão pelas valiosas trocas de conhecimento, reflexões e exemplos que ampliaram minha compreensão sobre o papel da gestão e da responsabilidade na construção de um futuro mais eficiente, ético e humano. Especialmente ao meu Orientador: Gustavo Correia Xavier, que participou de boa parte de minha trajetória na graduação (JGP, Monitoria, Liga Financeira – UFPB3, Estágio, entre outros projetos), contribuindo de forma marcante para minha Formação Acadêmica.

Sou grato também aos desafios que testaram meus limites, exigiram maturidade e ensinaram o verdadeiro valor da persistência. Cada obstáculo superado contribuiu para moldar não apenas o profissional, mas também a pessoa que me tornei: mais consciente, analítico e comprometido em transformar conhecimento em impacto real.

Por fim, agradeço a todos que, de forma direta ou indireta, fizeram parte dessa trajetória. Este trabalho reflete não apenas meu esforço individual, mas também as influências, experiências e pessoas que contribuíram para minha caminhada.

A todos vocês, minhas sincera gratidão.

"E se olhares por muito tempo para o abismo, o abismo também olha para dentro de você." (Nietzsche, 1886)

RESUMO

Este Trabalho de Conclusão de Curso tem como objetivo analisar o desempenho do modelo econométrico clássico ARIMA (Box e Jenkins, 1976) e do modelo de aprendizado profundo TimesNet (WU et al., 2022), aplicados à previsão de séries temporais financeiras, com atenção à relação entre sinal e ruído, elemento central que define o grau de previsibilidade e a natureza linear ou não linear dos dados. A pesquisa avaliou a eficácia de cada abordagem em termos de acurácia estatística. A metodologia incluiu o uso de métricas amplamente reconhecidas, como RMSE, MAE e MAPE, aplicadas a ativos representativos de economias maduras e emergentes, incluindo índices de mercado acionário (americano, brasileiro, alemão, Reino Unido e Japão) e a commodity ouro. Os resultados empíricos mostraram que o modelo ARIMA, baseado em uma estrutura linear e parcimoniosa, teve desempenho superior em horizontes temporais curtos e séries com baixo sinal e alto ruído, confirmando sua eficiência em ambientes com menor previsibilidade. Por outro lado, o TimesNet, ao empregar mecanismos de aprendizado profundo capazes de capturar padrões não lineares e dependências complexas, apresentou vantagens apenas quando combinado com o ARIMA, reforçando o potencial dos modelos híbridos. Assim, conclui-se que a integração entre modelos clássicos e técnicas avançadas de inteligência artificial permite explorar de forma complementar sinais lineares e não lineares, aprimorando a previsão de retornos e contribuindo para decisões mais assertivas em um ambiente dinâmico na gestão de portfólios de investimentos.

Palavras-chave: Séries Temporais; ARIMA; Inteligência Artificial; Deep Learning; TimesNet; Mercado Financeiro.

ABSTRACT

This Undergraduate Thesis aims to analyze the performance of the classical econometric model ARIMA (Box and Jenkins, 1976) and the deep learning model TimesNet (Wu et al., 2022), applied to the forecasting of financial time series, with attention to the relationship between signal and noise, a central element that defines the degree of predictability and the linear or nonlinear nature of the data. The research evaluated the effectiveness of each approach in terms of statistical accuracy. The methodology included the use of widely recognized metrics, such as RMSE, MAE, and MAPE, applied to assets representing mature and emerging economies, including stock market indices (United States, Brazil, Germany, the United Kingdom, and Japan) and the gold commodity. Empirical results showed that the ARIMA model, based on a linear and parsimonious structure, performed better in short-term horizons and in series with low signal and high noise, confirming its efficiency in environments with lower predictability. On the other hand, TimesNet, by employing deep learning mechanisms capable of capturing nonlinear patterns and complex dependencies, presented advantages only when combined with ARIMA, reinforcing the potential of hybrid models. Thus, it is concluded that the integration between classical models and advanced artificial intelligence techniques allows for the complementary exploration of linear and nonlinear signals, enhancing return forecasting and contributing to more assertive decision-making in the dynamic context of investment portfolio management.

Keywords: Time Series; ARIMA; Artificial Intelligence; Deep Learning; TimesNet; Financial Market.

SUMÁRIO

1 INTRODUÇÃO	10
1.1. Fundamentos Técnicos.....	11
2 METODOLOGIA	15
2.1 Coleta e Preparação dos Dados.....	15
2.2 Arquitetura e Implementação dos Modelos.....	17
2.2.1. Arima recursivo onde se tem um grid Search.....	17
2.2.2 Modelo de Aprendizado Profundo: TimesNet.....	18
2.2.3 Modelo Híbrido (ARIMA combinado com TimesNet)	19
2.3 Métricas de Avaliação de Desempenho e Acurácia Estatística.....	21
2.4 Síntese do Framework Metodológico.....	22
3 RESULTADOS	23
3.1. Apresentação Consolidada dos Erros Preditivos.....	23
3.2. Análise por Horizonte Temporal.....	24
3.3. Análise por Contexto Econômico.....	26
3.3.1. Mercados Maduros (GSPC, FTSE, GDAXI).....	27
3.3.2. Ibovespa (BVSP) – O Desafio da Volatilidade Emergente.....	31
3.3.3. Mercado Asiático: Nikkei 255 (N225).....	33
3.3.4. Ouro (GC=F) – Dinâmica de Porto Seguro.....	34
3.4. Discussão dos resultados Obtidos.....	35
4 CONSIDERAÇÕES FINAIS	36
REFERÊNCIAS	38

1. INTRODUÇÃO

A previsão de retorno e risco de ativos financeiros emerge como um pilar central na literatura de finanças, com implicações diretas na gestão de portfólios e na avaliação de desempenho de fundos de investimentos. Tradicionalmente, a análise financeira baseia-se na Hipótese de Mercados Eficientes (HME), que, em sua formulação mais estrita, postula a impossibilidade de prever retornos consistentemente (Fama, 1998). Contudo, a crescente complexidade e volatilidade dos mercados globais, aliadas ao aumento exponencial do volume de dados e à maior capacidade de absorção de informações, impulsionam a demanda por modelos preditivos mais sofisticados para apoiar a tomada de decisão dos investidores (Liu et al., 2019; Mehtab, Sen & Dutta, 2020). Adicionalmente, a capacidade preditiva de modelos de séries temporais podem demonstrar leves desvios à HME em decorrência da existência de limites de arbitragem (Kellly et al., 2022).

O Aprendizado de Máquina (AM) consiste em um subconjunto da Inteligência Artificial voltado para o desenvolvimento de algoritmos capazes de aprender padrões a partir de dados, sem depender de instruções explícitas para cada tarefa, o aprendizado ocorre quando o sistema melhora seu desempenho em uma tarefa com base na experiência obtida, o que o diferencia dos métodos puramente determinísticos. Em aplicações financeiras e de séries temporais, o aprendizado de máquina busca identificar relações complexas entre variáveis como retornos, volatilidade e volume que são difíceis de capturar por modelos estatísticos lineares tradicionais (Kelly et al., 2022; Provost; Fawcett, 2013).

O Deep Learning (DL), por sua vez, representa uma evolução do aprendizado de máquina, fundamentada em estruturas com múltiplas camadas de processamento hierárquico. Conforme destacado por Masini, Medeiros e Mendes (2023), os modelos de aprendizado profundo permitem capturar relações não lineares complexas e dependências temporais de longo alcance, especialmente úteis em séries financeiras. Essa abordagem utiliza redes neurais profundas para decompor os dados em representações progressivamente mais abstratas, aprimorando a generalização e a capacidade de adaptação a diferentes regimes de mercado.

As Redes Neurais Artificiais (RNAs) constituem a base do aprendizado profundo e foram inspiradas na estrutura biológica do cérebro humano. De acordo com Haykin (2001), essas redes são compostas por unidades de processamento interconectadas (neurônios artificiais) que ajustam pesos sinápticos para minimizar erros de previsão, permitindo a aproximação de funções não lineares de alta complexidade. Estudos como o de Mehtab, Sen e Dutta (2020) reforçam essa vantagem ao demonstrarem que arquiteturas neurais, em capturar de forma mais eficiente as variações estruturais e de tendência dos mercados.

A comparação entre abordagens clássicas, como o ARIMA (Autoregressive Integrated Moving Average), proposto por Box e Jenkins (1976), e técnicas de Deep Learning, como as redes LSTM (Long Short-Term Memory), é amplamente estudada (Athey e Ibens, 2019). Outros estudos corroboram essa perspectiva, mostrando que modelos LSTM superam o ARIMA em setores como óleo e gás e na previsão de preços de ações, com a vantagem adicional de exigirem menos dados para previsões eficientes devido à sua alta capacidade de capturar o comportamento observado (Agrawal et al., 2023; Bezerra e Albuquerque, 2019).

Sendo um modelo linear clássico, o ARIMA consolidaram-se como referências na modelagem de séries temporais, na qual combina três componentes: autorregressão (AR), diferenciação (I) e médias móveis (MA). Destacando por robustez em estruturas lineares e estacionárias, sendo usado por décadas como padrão para avaliação de previsões financeiras. Ele permite quantificar a previsibilidade dos retornos financeiros explicada pela autocorrelação linear, fornecendo uma base sólida e comparável para avaliar o desempenho de modelos mais

complexos. (Box e Jenkins, 1976; Tavares; Quadrelli, 2022).

A ascensão do aprendizado de máquina (AM) consolidou-se como uma técnica proeminente na ciência de dados, para prever os preços de mercado e subsidiar decisões de investimento é essencial. Redes Neurais Artificiais, particularmente as redes LSTM, têm demonstrado eficiência na previsão de preços de mercado. Estudos como os de Liu et al. (2019) e Mehtab, Sen e Dutta (2020) indicam que essa técnica tem o poder de analisar grandes volumes de dados históricos, identificar padrões e tendências, e prever com razoável precisão os preços futuros de ativos de renda variável, permitindo decisões mais informadas.

Proposta por Hochreiter e Schmidhuber (1997), a LSTM é um tipo especial de Rede Neural Recorrente (RNN) projetada para resolver o problema da dissipação do gradiente, permitindo o aprendizado de dependências temporais de longo prazo em sequências de dados. O funcionamento das LSTM baseia-se em três estruturas chamadas "portões" (gates): o portão de esquecimento, que elimina informações irrelevantes; o portão de entrada, que adiciona novas informações ao estado da célula; e o portão de saída, que determina as informações a serem transmitidas para a próxima previsão.

O TimesNet, em particular, representa um avanço relevante ao incorporar uma abordagem de convoluções temporais bidimensionais capazes de identificar múltiplos ciclos sazonais semanais, mensais ou anuais dentro de uma mesma série temporal. Diferentemente das RNNs ou LSTMs, que processam dependências temporais de forma sequencial, o "TimesNet" utiliza uma abordagem baseada em convoluções temporais 2D, permitindo maior eficiência na aprendizagem de padrões complexos e sazonais, mantendo a hipótese de coexistência de diferentes regimes de ao mesmo tempo (Wu et al., 2023).

Nesse cenário, o campo das finanças quantitativas busca modelos robustos capazes de lidar com a alta dimensionalidade e a incerteza inerente aos ativos de risco. Este estudo propõe uma análise comparativa entre duas metodologias fundamentais na previsão de séries temporais: o modelo estatístico clássico ARIMA e a rede neural recorrente avançada LSTM, investigando qual abordagem oferece maior acurácia preditiva em ambientes financeiros complexos de curto prazo (Gu, Kelly & Xiu, 2020; Tsay, 2010; Kelly et al., 2024).

1.1. Fundamentos Técnicos

A busca por métodos preditivos precisos e operacionalmente viáveis, cujo investimento seja inferior ao custo de implementação prática no mercado real, capazes de apoiar a inteligência estratégica em finanças (Tsay, 2010; Gu, Kelly & Xiu, 2020), justifica a comparação detalhada desses dois modelos. A análise da eficácia desses modelos para identificar momentos e antecipar movimentos de mercado fortalecerá o uso de métodos quantitativos no suporte à tomada de decisão em investimentos.

Giannone et al. (2021) analisaram aplicações em finanças, microeconomia e macroeconomia, avaliando a premissa de modelos econômicos de baixa dimensão e encontrando suporte limitado para essa hipótese. Gu et al. (2020) evidenciaram outra aplicação convincente de ML, utilizando dados de alta dimensão para prever retornos do mercado de ações. Goulet Coulombe et al. (2022) demonstraram que a capacidade dos modelos de ML em lidar com não linearidades melhora significativamente as previsões macroeconômicas.

A otimização de Hiperparâmetros define o controle do modelo e, antes do treinamento, representa variáveis de controle que influenciam diretamente a capacidade de generalização e desempenho de um modelo de aprendizado de máquina. Diferentemente dos parâmetros ajustados durante o aprendizado, os hiperparâmetros configuram o comportamento do

algoritmo, como o número de camadas e neurônios em uma rede neural, a taxa de aprendizado (learning rate) ou o tamanho da janela temporal em séries financeiras. Esses ajustes equilibram a complexidade e estabilidade do modelo, sendo essenciais para que ele capture de forma eficiente as estruturas temporais e dinâmicas não lineares do mercado financeiro (Haykin, 2001; Mehtab; Sen; Dutta, 2020).

Provost e Fawcett (2013) ressaltam que a definição apropriada das variáveis de entrada é uma das etapas mais críticas na modelagem preditiva, pois impacta diretamente a capacidade do modelo de extrair conhecimento útil e representar as relações subjacentes nos dados. Os dados fornecidos ao modelo (inputs) correspondem às entradas, enquanto as features representam variáveis explicativas ou atributos derivados que descrevem o comportamento da série temporal, como retornos, volatilidade, médias móveis ou outros indicadores relevantes para o aprendizado. Gu, Kelly e Xiu (2020) enfatizam que a seleção e transformação de features em finanças quantitativas devem preservar a integridade informacional e refletir fielmente a estrutura estocástica das séries, assegurando previsões consistentes e empiricamente robustas.

Tsay (2010) discute que, em séries financeiras, a relação entre sinal e ruído determina o quanto os movimentos observados refletem informações relevantes (sinal) em oposição a flutuações aleatórias (ruído). O sinal representa a componente informativa que os modelos buscam identificar para prever tendências, enquanto o ruído corresponde à parte imprevisível e aleatória dos preços, característica de mercados emergentes. A predominância do ruído dificulta a identificação de padrões consistentes, tornando modelos lineares, como o ARIMA, mais adequados em contextos de baixo sinal e alta variabilidade. Por outro lado, em ambientes com dependências complexas e persistentes, modelos como o TimesNet, tendem à capturar estruturas subjacentes que escapam às abordagens tradicionais (Mehtab, Sen e Dutta, 2020).

As aplicações de teste preditivo buscam evitar vazamentos de dados (data leakage), que ocorrem quando informações do conjunto de teste, que deveriam permanecer desconhecidas durante o treinamento, são utilizadas direta ou indiretamente pelo modelo no processo de ajuste. Essa prática gera um viés otimista nos resultados e compromete a validade da avaliação preditiva. De acordo com Provost e Fawcett (2013), a separação entre treino, validação e teste é um princípio fundamental da metodologia de ciência de dados, garantindo que o modelo seja avaliado sob condições realistas, simulando previsões em dados futuros.

Essas aplicações também passam por um processo de normalização, essencial em modelos baseados em redes neurais artificiais, já que essas estruturas são sensíveis à escala das variáveis de entrada. A normalização ajusta os dados para um intervalo, permitindo que todas as features (características explicativas captadas pelo modelo) contribuam de forma equilibrada para o aprendizado, evitando que variáveis de maior magnitude dominem o processo de ajuste dos pesos sinápticos, melhorando a estabilidade e convergência (Haykin, 2001).

Kotchoni et al. (2019) investigaram seis categorias de modelos para prever diferentes séries, comparando a eficácia preditiva entre modelos tradicionais e de aprendizado de máquina (ML). Métodos mais recentes, como mineração de dados e análise de sentimento, também trouxeram avanços à área, como demonstrado por Duarte et al. (2021), que utilizaram notícias para prever preços de ações. De forma semelhante, Saurabh e Dey (2020) apresentaram vários modelos preditivos baseados em ML para analisar a relação entre preços de ações e reações do mercado capturadas por tweets relevantes, reforçando a importância de aplicar múltiplos modelos para prever ativos financeiros.

Araujo e Gaglianone (2023) compararam 50 métodos de previsão para a inflação brasileira, concluindo que os modelos de ML superam frequentemente os métodos tradicionais (como o modelo ARIMA). Contudo, eles também destacaram a ausência de um método

universalmente superior, com a escolha ideal dependendo dos dados e da métrica de avaliação. Garcia et al. (2017) aplicaram modelos de alta dimensão para prever a inflação brasileira, observando que tais modelos, especialmente a regressão de subconjuntos completos, apresentaram um desempenho excepcional em diversos horizontes de previsão. Na previsão do mercado de ações utilizando ML, Pierdzioch e Risso (2018) testaram a hipótese das expectativas racionais ao preverem o mercado de ações americano com árvores de regressão impulsionadas, comparando os resultados em diferentes janelas temporais.

Ersan et al. (2020) compararam o desempenho de três métodos para prever a direção de séries temporais financeiras, demonstrando que o aumento do tamanho da janela de tempo é benéfico somente para dados horários e até certo ponto. Kotchoni et al. (2019) concluíram que o modelo ARIMA é eficaz para prever mudanças na inflação no curto prazo, enquanto a Média de Modelos Rica em Dados Regularizada (RDRMA) se destaca em horizontes mais longos.

Diversos estudos na literatura corroboram o desempenho superior do modelo ARIMA em horizontes de curto prazo (até 5 dias), especialmente em ambientes caracterizados por baixo sinal informacional e alta volatilidade. De acordo com Tsay (2010), modelos lineares, ao representarem adequadamente a estrutura autorregressiva e estacionária das séries financeiras, tendem a produzir previsões mais estáveis sob ruído elevado, situação em que as relações de dependência são predominantemente locais. Essa robustez em curtos horizontes é reforçada por Faust e Wright (2013), que argumentam que, em mercados próximos da eficiência informacional, a presença de ruído e a ausência de padrões persistentes reduzem o potencial de ganho de modelos complexos.

Por outro lado, Mehtab, Sen e Dutta (2020) mostram que redes LSTM (Long Short-Term Memory) apresentam maior eficiência em horizontes longos, onde há maior disponibilidade de dependências temporais e padrões estruturais a serem aprendidos. No entanto, em períodos curtos ou intradiários, essas redes tendem a sofrer com sobreajuste (overfitting) e degradação de desempenho, dada sua sensibilidade a ruídos e flutuações de alta frequência. Em consonância, Gu, Kelly e Xiu (2020) evidenciam que o poder explicativo de modelos baseados em aprendizado de máquina depende fortemente da dimensionalidade e do horizonte de previsão, sendo os modelos lineares mais eficientes em contextos de ruído elevado e menor persistência informacional.

Entretanto, vale se destacar que os LSTM representam um avanço significativo no campo do Deep Learning e da ciência de dados, com uma estrutura permite que as LSTM mantenham e atualizem informações relevantes ao longo do tempo, oferecendo maior flexibilidade na captura de padrões não-lineares e de longa duração em séries financeiras (Gu; Kelly; Xiu, 2020; Masini; Medeiros; Mendes, 2023).

Um fator frequentemente apontado como determinante para o desempenho dos modelos é o horizonte de tempo. A literatura especializada evidencia que o desempenho relativo dos modelos de previsão é altamente sensível ao domínio de aplicação e ao horizonte temporal considerado. Por isso, a superioridade de técnicas como ARIMA ou LSTM não é universal, sendo essencial uma avaliação contextualizada. Em um estudo sobre previsão de curto prazo da demanda de potência ativa, o modelo ARIMA alcançou 46% de eficiência, enquanto o LSTM obteve somente 13% de acerto, considerado um espaço de previsão de curto prazo para ambos os modelos (Tavares & Quadrelli, 2022).

Ersan et al. (2020), Kotchoni et al. (2019) e Faust e Wright (2013), reforçam que a literatura ainda apresenta resultados ambíguos quanto à superioridade sistemática de um modelo superior aos demais. Enquanto o ARIMA tende a oferecer previsões mais consistentes em horizontes curtos (5 dias) e dados relativamente estáveis, o LSTM demonstra maior

desempenho em horizontes longos (15 dias) e em séries caracterizadas por não linearidade e alta volatilidade (Tsay, 2010; Gu, Kelly e Xiu, 2020).

Como apresentado no estudo de Gonçalves, Alexandre e Lima (2023), foi realizada uma análise comparativa entre os modelos ARIMA e LSTM aplicada a séries temporais financeiras, com o objetivo de avaliar o desempenho de cada abordagem em diferentes janelas de previsão especificamente de 5, 15 e 30 dias. Os autores utilizaram métricas clássicas de erro, como o RMSE (Root Mean Square Error) e o MAE (Mean Absolute Error), para mensurar a precisão dos modelos. A tabela a seguir, apresenta um resumo dos resultados obtidos por Gonçalves, Alexandre e Lima (2023), destacando a superioridade do modelo ARIMA em curtos prazos e a eficiência do LSTM em janelas mais amplas:

Tabela 1: Desempenho comparativo entre ARIMA e LSTM em diferentes horizontes temporais

Horizonte/dias	Modelo	MAPE	RMSE
5 dias	ARIMA	0,0108	223,24
	LSTM	0,0401	741,24
15 dias	ARIMA	0,0367	777,54
	LSTM	0,0167	438,81
30 dias	ARIMA	0,0616	1.295,37
	LSTM	0,0141	359,3
Série completa	ARIMA	0,0527	1.191,97
	LSTM	0,0251	581,06

FONTE: Adaptado de Gonçalves, Alexandre e Lima (2023).

Essa dualidade sugere que a escolha do modelo deve ser condicionada ao horizonte de previsão e à natureza da série temporal em análise, tema que permanece como uma lacuna central no campo da econometria financeira, podendo afetar significativamente o desempenho do modelo (Keelawat et al., 2021).

Tavares e Quadrelli (2022) realizaram uma pesquisa sobre previsão de demanda de potência ativa, utilizando os modelos ARIMA e LSTM. Nesse estudo, o ARIMA apresentou 46% de precisão, enquanto o LSTM obteve apenas 13%, indicando que o desempenho de cada modelo depende das características das séries temporais fornecidas, um consenso universal na literatura. Onde pesquisas, concluíram que o modelo preditivo mais eficaz varia conforme o horizonte considerado.

Gu, Kelly e Xiu (2020), trazem a perspectiva de que o uso de janelas deslizantes (Rolling Out ou Rolling Deslizante) é essencial em finanças empíricas porque os relacionamentos entre variáveis mudam com o tempo, e modelos fixos podem rapidamente se tornar obsoletos diante de novas dinâmicas de volatilidade e regimes de mercado. Assim, o Rolling Out possibilita a adaptação contínua dos parâmetros do modelo, reduzindo o risco de sobreajuste e aumentando a relevância das previsões, sendo o mais realista para um mercado altamente dinâmico.

Na mesma direção, Masini, Medeiros e Mendes (2023) destacam que estratégias baseadas em janelas móveis são fundamentais para modelos de aprendizado de máquina aplicados a finanças, pois permitem o reequilíbrio periódico entre dados recentes e históricos, garantindo que o modelo aprenda de maneira dinâmica em ambientes não estacionários. Consequentemente, o uso do out deslizante neste trabalho visa reproduzir um cenário preditivo

mais realista, onde o modelo é constantemente atualizado com informações recentes.

Diante das evidências acima, surge a necessidade de um estudo sobre modelos híbridos para superar essas limitações, combinando tanto a capacidade linear de curto prazo do ARIMA quanto as do modelo TimesNet de Wu et al. (2023), resultando na integração de ambos. O ARIMA tem demonstrado eficácia no curto prazo, enquanto o TimesNet se apresenta na identificação de sazonalidades, posicionando o TimesNet como uma alternativa promissora com características inovadoras, apesar de haver poucos estudos que aprofundem sua aplicação no mercado financeiro.

A metodologia, fundamentada em uma abordagem quantitativa e comparativa, foi desenvolvida com etapas bem definidas. Após a coleta e preparação minuciosa dos dados, implementam-se os modelos preditivos, sendo o ARIMA escolhido por sua consolidação na literatura de séries temporais, e o TimesNet, pela capacidade de lidar com dinâmicas temporais complexas. Em seguida, realiza-se a otimização de parâmetros de ambos os modelos, culminando na avaliação de desempenho com ênfase no modelos com menor RMSE (Raiz Quadrada do Erro-médio)

2. METODOLOGIA

O presente estudo adota uma abordagem quantitativa e comparativa para avaliar a eficácia de modelos econométricos e de aprendizado profundo na previsão de séries temporais financeiras. A pesquisa analisa e compara modelos econométricos clássicos e de aprendizado profundo, utilizando uma estrutura metodológica cuidadosamente elaborada em etapas sequenciais. Essas etapas incluem a coleta e preparação de dados, a implementação de dois modelos distintos ARIMA e TimesNet, como modelo avançado, a otimização de seus parâmetros e a avaliação de desempenho com um conjunto robusto de métricas. Cada decisão no processo foi planejada para assegurar rigor científico, reproduzibilidade dos resultados e relevância prática no contexto dos mercados financeiros.

2.1 Coleta e Preparação dos Dados

A qualidade e a integridade dos dados são fatores cruciais e indispensáveis para a validade e robustez de qualquer modelo preditivo financeiro, considerando a natureza intrinsecamente complexa, volátil e ruidosa desses mercados. As fontes de dados históricos de preços foram sistematicamente coletadas utilizando a biblioteca yfinance, uma ferramenta amplamente reconhecida e utilizada no meio acadêmico e de pesquisa para o acesso a dados de mercado em larga escala (Neto, 2010; Goodell et al., 2021).

A seleção de múltiplos ativos, incluindo índices globais e commodities, e a análise em diferentes granularidades (diária e intradiária) seguem a abordagem de estudos que testam a robustez e generalização de modelos preditivos em diferentes condições de mercado. Trabalhos como o de Ersan, Nishioka e Scherp (2020) demonstram que a eficácia de um modelo de previsão pode variar significativamente conforme a frequência dos dados e o nível de volatilidade do ativo. Assim, a inclusão de horizontes distintos permite avaliar se os modelos mantêm desempenho consistente tanto em escalas curtas (horárias) quanto longas (diárias), testando a adaptabilidade das técnicas a diferentes regimes temporais e dinâmicas de ruído.

A ideia central dessa seleção é aplicar os algoritmos sobre um conjunto diversificado de economias e classes de ativos, abrangendo mercados maduros, emergentes e de commodities, a fim de verificar o poder de generalização e estabilidade dos modelos. Essa diversidade é essencial, pois o comportamento dos retornos e a estrutura de dependência entre observações

variam substancialmente conforme o grau de desenvolvimento econômico e a profundidade do mercado (Gu, Kelly & Xiu, 2020; Tsay, 2010; Kelly et al., 2024).

Para as análises em diferentes escalas temporais, foi selecionada uma amostra de 2.500 dias para o horizonte diário, representando aproximadamente dez anos de dados, período considerado suficiente para capturar diferentes ciclos econômicos e regimes de volatilidade (Tsay, 2010). A amostra de 502 dias permite avaliar previsões em uma janela intermediária de cerca de dois anos, adequada à comparação entre curto e médio prazo. Por fim, a frequência intradiária de 1.752 horas (aproximadamente 251 dias úteis), esses intervalos seguem o padrão da literatura em séries financeiras de alta frequência (Tsay, 2010; Hyndman & Athanasopoulos, 2018), permitindo equilibrar tamanho amostral e estabilidade de estimativas.

Essa combinação temporal equilibra extensão histórica e granularidade operacional, possibilitando avaliar se os modelos mantêm estabilidade e precisão sob diferentes escalas de observação e níveis de volatilidade um aspecto crucial na literatura de previsão de séries financeiras (Gu, Kelly & Xiu, 2020). Servindo como base para observar o desempenho dos modelos em ambientes de alta frequência e baixo sinal-ruído, contexto no qual a capacidade de generalização é frequentemente testada (Ersan, Nishioka & Scherp, 2020).

No contexto deste estudo, foram incluídos índices como o americano, representado pelo S&P 500 (GSPC), refletindo uma economia desenvolvida e altamente líquida; o brasileiro, representado pelo Ibovespa (BVSP), que reflete um mercado emergente, sujeito a maior volatilidade e choques regionais; o japonês, representado pelo Nikkei 225 (N225), indicando uma economia asiática madura com forte influência tecnológica; o alemão, representado pelo DAX (GDAXI), expressando o comportamento de uma das principais economias industriais da Europa; e o britânico, representado pelo FTSE 100 (FTSE), que reúne as maiores empresas do Reino Unido e é referência para mercados europeus. A commodity ouro (GC=F) foi incluída como ativo de refúgio (safe haven) e commodity global, com comportamento de preço frequentemente inverso aos índices acionários em períodos de incerteza.

Como se pode observar, os índices escolhidos são os mais representativos de suas economias. Essas escolhas possibilitam testar se a capacidade preditiva dos modelos se mantém em diferentes estruturas de mercado desde economias altamente integradas e eficientes até mercados em desenvolvimento, com maiores assimetrias informacionais e comportamentos menos previsíveis, conforme discutido por Faust e Wright (2013) e Ersan, Nishioka e Scherp (2020). Tais estudos reforçam que o grau de eficiência de mercado e o nível de ruído nas séries temporais influenciam diretamente o desempenho dos modelos de previsão, corroborando os pressupostos da Hipótese de Mercado Eficiente (HME) e a importância de testar modelos sob diferentes regimes econômicos e granularidades de dados.

Eventuais valores ausentes nos conjuntos de dados foram tratados e imputados utilizando a técnica de forward fill (ffill), que consiste em preencher as lacunas com o último valor válido observado, como sugerido por, Mehtab, Sen e Dutta (2020). Onde são amplamente utilizadas em contextos de alta frequência, por preservarem a ordem cronológica e a integridade informacional das séries temporais financeiras. Conforme destacado por Tsay (2010), a manutenção da característica de caminho aleatório e da dependência temporal é essencial para garantir a validade das análises e previsões em séries financeiras. Além disso, Gu, Kelly e Xiu (2020) reforçam que o uso de dados consistentes e o controle de distorções no pré-processamento são fatores determinantes para a robustez empírica de modelos preditivos.

O conjunto de dados foi dividido em amostras de treino, validação e teste. A amostra de treino é utilizada para o aprendizado dos parâmetros do modelo; a de validação, para o ajuste fino dos hiperparâmetros; e a de teste, mantida "invisível" durante todo o processo, é reservada

exclusivamente para a avaliação final do desempenho, assegurando resultados imparciais. Sendo aplicada uma divisão de 80% para treinamento, 10% para validação e os 10% restantes para teste (Provost; Fawcett, 2013)

O conjunto de validação é utilizado especificamente para otimizar o desempenho dos modelos, permitindo ajustes refinados sem comprometer a imparcialidade. A performance final e a capacidade de generalização dos modelos são avaliadas exclusivamente no conjunto de teste, que permanece intocado durante as fases de treinamento e validação. Essa separação é de suma importância para garantir a credibilidade e a imparcialidade em estudos de aprendizado de máquina aplicados à precificação de ativos e previsão financeira (Gu; Kelly; Xiu, 2020).

No presente estudo, os dados foram submetidos ao processo de normalização, no qual o escalonador MinMaxScaler foi ajustado com base no conjunto de treinamento, garantindo que nenhuma informação do conjunto de teste influenciasse os parâmetros do modelo. Isso assegura que os features contribuam de forma equilibrada para o modelo e melhora a estabilidade do treinamento por gradiente descendente (Haykin, 2001).

Para evitar vazamento de dados (data leakage), o escalonador foi configurado exclusivamente com os dados do conjunto de treino, uma prática essencial na metodologia de ciência de dados (Provost; Fawcett, 2013). Além disso, a normalização dos dados facilita a captura de padrões sutis e relações não lineares entre variáveis, permitindo que o modelo identifique estruturas temporais e dependências complexas sem distorções causadas por escalas incompatíveis, ou volatidades do mercado.

2.2 Arquitetura e Implementação dos Modelos

2.2.1. Arima recursivo onde se tem um grid search

Essa metodologia é amplamente aplicada na modelagem de séries temporais, combinando componentes autorregressivos e de média móvel em um único framework, lidando tanto com padrões lineares quanto com a não estacionalidade por meio da diferenciação. A escolha do modelo ARIMA como benchmark é justificada por seu status canônico na literatura de séries temporais. Sua eficácia em modelar estruturas lineares o torna um ponto de partida robusto e um concorrente desafiador para modelos mais complexos, sendo até hoje amplamente utilizado em estudos comparativos (Tavares; Quadrelli, 2022; Gonçalves; Alexandre; Lima, 2023).

A otimização dos parâmetros p, d e q do modelo ARIMA foi realizada por meio de uma grid search, permitindo uma exploração sistemática do espaço de modelos. Essa abordagem garante total transparência sobre os parâmetros testados e as configurações que levaram aos melhores resultados. Esse processo de otimização envolve uma avaliação exaustiva de combinações específicas para minimizar o erro de previsão, sendo uma alternativa robusta aos métodos automáticos, tendo como critério para seleção os parâmetros com menor RMSE (Hyndman; Athanasopoulos, 2018).

O modelo ARIMA(p, d, q) combina três componentes principais: Autorregressivo (AR(p)), onde o valor atual da série é explicado por uma combinação linear de valores passados; Integração (I(d)), que indica o número de diferenciações necessárias para tornar a série estacionária; e Média Móvel (MA(q)), em que o valor da série é influenciado por erros passados, conforme descrito por Box e Jenkins (1976). Esse processo possibilita a adaptação do modelo a diferentes ativos e economias, ajustando-se ao comportamento da série histórica para projetar valores futuros.

Adotou-se uma estratégia de previsão recursiva, na qual o modelo é reestimado a cada

novo passo temporal, simulando um ambiente real em que novas informações se tornam disponíveis, alinhando-se às práticas de previsão em tempo real. Sendo uma estratégia uma das maneiras mais fiel um cenário de implantação em tempo real. Nessa abordagem, o modelo é periodicamente reestimado com a inclusão de novas informações, permitindo sua adaptação às mudanças nas dinâmicas da série temporal (Garcia; Medeiros; Vasconcelos, 2017).

A validação do modelo ARIMA não se restringiu à sua acurácia preditiva. Realizou-se uma análise de diagnóstico dos resíduos para verificar se eles se comportavam como ruído branco, uma premissa essencial que indica que o modelo capturou toda a estrutura de dependência temporal presente nos dados, conforme a metodologia clássica. A ausência de autocorrelação estatisticamente significativa nos resíduos é um indicador fundamental de que o modelo conseguiu capturar a estrutura linear preditiva dos dados, deixando apenas o "ruído" aleatório não explicável (Montgomery et al., 2008).

No presente estudo, o ARIMA foi implementado, utilizando informações desde o período de treinamento até o instante-alvo para prever o próximo valor da série. Para previsões intradiárias, o modelo considerou a última hora disponível para antecipar o movimento da próxima hora. Já para previsões de horizonte diário, utilizou-se aproximadamente 80% para treinamento, 10% para validação, reservando 10% o teste na amostra. Essa abordagem permite o uso do ARIMA medir a eficácia de modelos mais sofisticados (Box; Jenkins, 1976; Montgomery et al., 2008).

2.2.2 Modelo de Aprendizado Profundo: TimesNet

A literatura recente destaca consistentemente a superioridade desses modelos em previsões de séries temporais financeiras, em longo prazo (períodos acima de 15 dias). Para abordar dinâmicas não-lineares, complexas e de longa dependência, inerentes às séries financeiras foi implementado um modelo de aprendizado profundo. Estudos recentes reforçam a eficácia superior desses modelos em diversas tarefas de precificação de ativos e previsão em longo prazo (Tsay, 2010; Masini; Medeiros; Mendes, 2023).

A Estrutura do Problema, a série temporal foi transformada em um formato de aprendizado supervisionado utilizando a técnica de janelas deslizantes. Essa abordagem, que gera pares de sequências de entrada e saída, é um método eficaz e amplamente utilizado para aplicar arquiteturas de aprendizado profundo a problemas de previsão, como demonstrado por Kim et al. (2020). A série temporal original foi estruturada nesse formato supervisionado por meio das janelas deslizantes, permitindo que o modelo mapeie sequências passadas de dados para prever valores futuros.

A arquitetura TimesNet, baseada em Redes Neurais Convolucionais 1D (Conv1d), foi projetada especificamente para funcionar como um extrator de características hierárquicas. Sua principal capacidade está em identificar padrões locais e recorrentes em diferentes escalas temporais, sendo especialmente útil para capturar interdependências complexas presentes nos dados de séries temporais, como apresentando por Wu, H. et al. (2023). Modelos híbridos que integram LSTM para extração de features têm apresentado resultados promissores na previsão de séries financeiras (Zhang et al., 2019; Shi et al., 2022).

Para a otimização dos hiperparâmetros do TimesNet, foi empregada a otimização Bayesiana com a biblioteca Optuna, visando calibrar a taxa de aprendizagem, tamanho do lote e filtros convolucionais, confere maior robustez ao modelo, minimizando problemas de sobreajuste (overfitting), no qual o modelo se ajusta excessivamente ao ruído presente na amostra de treino, comprometendo sua capacidade de generalização. Essa técnica destaca-se por sua eficiência e sistematicidade na busca de configurações ideais, considerando a

sensibilidade dos modelos de aprendizado profundo a esses parâmetros (Goodell et al., 2021).

O processo foi configurado para minimizar o erro na amostra de validação, evitando enviesamento pelos dados de teste, com um comprimento de entrada de 12 para as últimas observações do diário e 2 para o intradiário, permitindo que o modelo se adapte à diversidade dos dados (temporais e econômicos). O treinamento do modelo foi realizado com o otimizador Adam, conhecido por sua eficiência na gestão de taxas de aprendizado, proporcionando uma convergência mais rápida e estável durante o treinamento da rede neural. O uso do Adam é uma escolha robusta e amplamente validada para problemas de regressão com redes neurais. O treinamento final, após a otimização, utiliza os melhores hiperparâmetros para ajustar o modelo final, consolidando o aprendizado obtido na fase de validação (Mehtab; Sen; Dutta, 2020; Haykin, 2001).

Em resumo, a LSTM apresenta vantagens evidentes sobre modelos lineares, como o ARIMA, na identificação de relações complexas e não lineares em séries financeiras. No entanto, sua eficácia depende diretamente do horizonte de previsão considerado, demonstrando que o desempenho relativo entre modelos econôméticos tradicionais e de aprendizado profundo não é absoluto, mas varia conforme o contexto e a granularidade temporal da análise. Diante dessas limitações, utilizou-se o modelo TimesNet, introduzido por Wu, H. et al. (2023), projetado para superar as restrições dos modelos lineares na captura de dinâmicas não lineares e complexas, como a volatilidade em clusters e outros fenômenos autorregressivos não lineares frequentemente observados em séries financeiras.

2.2.3. Modelo Híbrido (ARIMA combinado com TimesNet)

A premissa de que um único modelo preditivo pode capturar toda a complexidade de uma série temporal financeira é frequentemente irrealista. Séries de preços de ativos são tipicamente compostas por diferentes estruturas, como tendências lineares, autocorrelação e padrões não-lineares complexos, que são mais bem modelados por diferentes classes de algoritmos. Diante dessa limitação, a implementação de estratégias de combinação (ensembles) e a criação de um modelo híbrido tornam-se fundamentais para esta análise (Tsay, 2010; Mehtab)

O objetivo é transcender as fraquezas de modelos individuais como a rigidez linear do ARIMA e a propensão ao overfitting do TimesNet, para construir uma previsão agregada que seja simultaneamente mais robusta, adaptativa e precisa, aproveitando as forças complementares de cada abordagem. A análise dos quatro modelos de combinação, Média Simples, Ponderação por Erro, Rolling Deslizante (Out-of-Sample) e Rolling Expansivo (In-Sample), revela distinções claras em termos de robustez e adaptabilidade (Zhang et al., 2019; Shi et al., 2022).

Os modelos de Média Simples e Ponderação por Erro representam as formas mais diretas de combinação de previsões, um conceito cuja eficácia foi estabelecida no trabalho seminal de Bates e Granger (1969). A Média Simples, ao atribuir pesos idênticos (0.5) ao ARIMA e ao TimesNet/LSTM, opera sob a premissa de que não há informação prévia suficiente para favorecer um modelo em detrimento do outro. Sua força reside na diversificação de erros, que tende a produzir uma previsão agregada com variância menor do que a dos modelos individuais.

A Ponderação por Erro refina essa abordagem ao atribuir pesos fixos, calculados com base no desempenho histórico (RMSE) em um conjunto de validação. Esta técnica busca dar maior relevância ao modelo que demonstrou maior acurácia no passado. A equação geral para a combinação pela média simples pode ser expressa como:

$$\hat{y}_t^{MS} = \frac{1}{2}(\hat{y}_t^{ARIMA} + \hat{y}_t^{DL})$$

A ponderação por erro, por sua vez, refina essa abordagem ao atribuir pesos calculados com base no desempenho histórico dos modelos, geralmente mensurado pelo RMSE em um conjunto de validação, tendo sua equação sendo expressa como:

$$\hat{y}_t^{PE} = w_1\hat{y}_t^{ARIMA} + w_2\hat{y}_t^{DL}$$

sujeito a $w_1 + w_2 = 1$, com:

$$w_i = \frac{\frac{1}{RMSE_i}}{\sum_{j=1}^2 \frac{1}{RMSE_j}}$$

Ambas as abordagens são "estáticas" porque os pesos de combinação, uma vez definidos, não se alteram ao longo do horizonte de previsão. Embora representem uma melhoria consistente sobre o modelo de pior desempenho (TimesNet/LSTM), sua natureza não-adaptativa os torna vulneráveis a mudanças de regime nos dados, um fenômeno comum em mercados financeiros (Tsay, 2010). Se a performance relativa dos modelos base se altera ao longo do tempo, um ensemble estático não consegue se ajustar a essa nova realidade.

O modelo Rolling Deslizante (Out) representa uma evolução significativa as combinações estáticas, incorporando princípios de stacking (Wolpert, 1992) em um framework dinâmico e adaptativo. Ao treinar um meta-modelo de regressão linear a cada passo de tempo, utilizando apenas as previsões dos modelos base em uma janela móvel de dados recentes, esta abordagem aprende continuamente a combinação ótima. A principal vantagem desta metodologia é sua capacidade de "esquecer" dados passados que podem não ser mais representativos da dinâmica atual do mercado. Sua formulação pode ser expressa como:

$$\hat{y}_t^{RD} = \alpha_t\hat{y}_t^{ARIMA} + \beta_t\hat{y}_t^{DL} + \varepsilon_t$$

em que os parâmetros α_t e β_t são reestimados a cada passo temporal, com base em uma janela móvel de tamanho w , de modo a minimizar o erro quadrático médio local:

$$\min_{\alpha_t, \beta_t} \sum_{i=t-w}^{t-1} (y_i - (\alpha_t\hat{y}_i^{ARIMA} + \beta_t\hat{y}_i^{DL}))^2$$

Essa capacidade de adaptação a "concept drift" mudanças na relação estatística subjacente dos dados explica seu desempenho superior em quase todos os cenários. Em

ambientes voláteis e não-estacionários, a relevância da informação decaiu com o tempo. O Rolling Deslizante captura essa realidade, permitindo que os pesos da combinação se ajustem dinamicamente para refletir qual dos modelos base (ARIMA ou TimesNet/LSTM) é mais confiável nas condições de mercado mais recentes. Sua eficácia empírica, tanto em mercados maduros quanto emergentes, válida a hipótese de que a adaptação contínua é crucial para a previsão financeira. Sendo particularmente eficaz em contextos não estacionários e de mudança de regime, fenômeno amplamente documentado na literatura financeira (Tsay, 2010).

Em contraste, o modelo de Rolling Expansivo utiliza uma janela de dados que cresce continuamente, incorporando todo o histórico disponível a cada passo de tempo. Embora também seja uma abordagem adaptativa, sua premissa fundamental é a de que todo o passado é relevante para o futuro. Os resultados do estudo demonstram que, para as séries financeiras analisadas, essa premissa é não apenas falha, mas perigosa. A formulação geral segue a mesma estrutura linear:

$$\hat{y}_t^{RE} = \alpha_t \hat{y}_t^{ARIMA} + \beta_t \hat{y}_t^{DL} + \varepsilon_t$$

O desempenho desse modelo, pode ser evidenciado uma severa sensibilidade a eventos estruturais passados. Ao nunca "esquecer", o modelo fica excessivamente ancorado a regimes de mercado antigos, como crises ou bolhas, resultando em um viés persistente que degrada sua capacidade preditiva no presente. Podendo ser mais propenso a overfitting ao histórico, onde a tentativa de usar mais dados resulta em um modelo menos generalizável e robusto, reforçando a conclusão de que, em finanças, a relevância da informação é frequentemente local e dependente do tempo (Gu; Kelly; Xiu, 2020; Masini; Medeiros; Mendes, 2023).

2.3 Métricas de Avaliação de Desempenho e Acurácia Estatística

A seleção de um conjunto diversificado de métricas é essencial para uma avaliação completa, pois cada métrica oferece uma perspectiva única sobre os erros de previsão. As métricas escolhidas (MAE, RMSE, MAPE) são amplamente utilizadas na literatura de previsão de séries temporais e permitem tanto a quantificação do erro em termos absolutos e percentuais quanto a avaliação da qualidade do ajuste do modelo (Hyndman; Athanasopoulos, 2018; Tsay, 2010).

O uso combinado dessas métricas fornece uma visão abrangente e confiável do desempenho comparativo dos modelos. A avaliação foi dividida em duas categorias principais: acurácia estatística, garantindo uma análise completa e multifacetada. A análise final dos modelos foi estruturada em duas camadas complementares. A primeira aborda a acurácia estatística, medindo a proximidade das previsões aos valores reais. Essa abordagem dual é crucial, pois um modelo estatisticamente preciso pode não ser necessariamente lucrativo ou aplicável na prática (Mesquita; Oliveira; Pereira, 2020).

Erros percentuais, como MAPE (Mean Absolute Percentage Error) e são métricos adimensionais ideais para comparar o desempenho de modelos entre ativos com escalas de preço distintas. A adoção de um portfólio diversificado de métricas estatísticas é amplamente recomendada para uma avaliação abrangente e equilibrada do desempenho preditivo dos modelos, permitindo a análise dos algoritmos em diferentes economias e frequências (Hyndman; Athanasopoulos, 2018).

2.4 Síntese Metodológica

Este capítulo apresentou a metodológico utilizada, rigoroso e multifacetado, desenvolvido para abordar de forma sistemática e científica as questões de pesquisa propostas. A pesquisa segue um protocolo disciplinado e exaustivo de coleta, tratamento e preparação de dados, fundamental para a construção de modelos robustos. Em seguida, realiza uma análise comparativa aprofundada entre um modelo linear clássico (ARIMA), que serve como referência, e uma arquitetura de aprendizado profundo (TimesNet), explorando suas respectivas capacidades preditivas. As premissas para a construção do modelo híbrido que combina tanto o ARIMA quanto o TimesNet e as métricas de mensuração de erros empregadas para distinção da eficácia dos modelos e suas combinações.

Todos os processos de otimização de parâmetros foram conduzidos de forma sistemática e transparente, assegurando a validade interna dos resultados. A metodologia culmina em uma abordagem de avaliação dual, que combina a precisão estatística das previsões com a relevância financeira e econômica, proporcionando uma visão holística do desempenho dos modelos, sem negligenciar as delimitações e limitações inerentes a qualquer estudo empírico. Esta estrutura metodológica cuidadosamente planejada garante que os resultados, a serem apresentados e discutidos no capítulo seguinte, sejam não apenas robustos e bem fundamentados, mas também adequadamente contextualizados e diretamente aplicáveis às questões de pesquisa propostas, contribuindo de maneira significativa para o campo das previsões financeiras.

Considera-se que os mercados financeiros, embora altamente eficientes, não seguem um "Passeio Aleatório" perfeito. Existem padrões e ineficiências, particularmente em horizontes de tempo mais curtos, que podem ser explorados por modelos quantitativos. Essa premissa desafia a forma forte da Hipótese de Mercados Eficientes e fundamenta toda a análise preditiva em finanças (Tsay, 2010). O estudo assume que os dados históricos de preços e volumes contêm informações relevantes sobre o comportamento futuro dos ativos. Modelos como o ARIMA baseiam-se explicitamente na autocorrelação, enquanto modelos de aprendizado profundo procuram capturar padrões não lineares mais complexos nos Mesmos Dados (Box; Jenkins, 1976; Masini Et Al., 2023).

O impacto do trabalho reside na sua contribuição para diferentes grupos, especialmente para a comunidade acadêmica e científica. A pesquisa aprimora o debate sobre a eficácia de modelos lineares versus não lineares em finanças, apresentando evidências empíricas atualizadas e uma metodologia transparente. Destaca-se por propor um modelo híbrido que combina ARIMA e TimesNet, buscando reduzir os erros de previsão. Embora as séries de preços sejam não estacionárias, considera-se que elas podem ser transformadas (e.g., por diferenciação ou análise de retornos) para apresentar propriedades estatísticas mais estáveis. No entanto, reconhece-se que a dinâmica dos mercados varia ao longo do tempo, justificando o uso de estratégias de validação robustas e modelos adaptativos (Hyndman; Athanasopoulos, 2018).

A qualidade do estudo e a confiabilidade das conclusões sobre relações de causa e efeito foram asseguradas por meio de comparações controladas entre os modelos, utilizando condições idênticas de dados e métricas. Para garantir que os resultados sejam generalizáveis além da amostra específica, os testes foram realizados em múltiplos ativos de mercados distintos (emergentes, desenvolvidos, commodities) e em diferentes frequências. Essa abordagem reforçou a validade externa. Ademais, a qualidade foi garantida pelo uso de um conjunto abrangente de métricas de avaliação que consideram tanto a precisão estatística quanto a relevância financeira, seguindo as recomendações de Hyndman e Athanasopoulos (2018), evitando inferências baseadas em um único indicador.

3. RESULTADOS

3.1. Apresentação Consolidada dos Erros Preditivos

A análise empírica inicia-se com a apresentação dos resultados quantitativos. A tabela a seguir consolida as métricas de desempenho dos modelos ARIMA, TimesNet e as Combinações aplicadas desses modelos em conjunto em todos os períodos e frequências avaliados. Esses dados constituem a base para a discussão teórica e as conclusões subsequentes, permitindo uma avaliação clara e objetiva do desempenho preditivo. A análise engloba a consolidação das métricas de erro desses modelos em diferentes períodos e mercados, fornecendo suporte à discussão posterior, (Athey, S.; Imbens, G. W., 2019). Por fim, apresentamos uma síntese dos achados, organizada por horizonte temporal e contexto econômico.

Tabela 2: Métricas de erros por frequência e Ativo

Índices	Modelo	Intradiário, 1752h			Diário, frequência de 502d			Diário, frequência de 2500d		
		RMSE	MAE	MAPE (%)	RMSE	MAE	MAPE (%)	RMSE	MAE	MAPE (%)
BVSP	ARIMA	442,93	272,93	0,19	1.196,10	891,6	0,64	1.257,80	960,02	0,72
	TimesNet (LSTM)	562,04	411,72	0,29	2.916,82	2.643,07	1,9	2.118,36	1.844,04	1,38
	Média simples	481,97	314,09	0,22	1.890,58	1.558,12	1,11	1.517,16	1.244,22	0,93
	Ponderada por erro	474,08	304,76	0,22	1.669,34	1.304,54	0,93	1.378,15	1.105,07	0,83
	Rolling deslizante (Out)	417,97	270,28	0,19	887,83	825,62	0,58	1.218,19	930,51	0,68
	Rolling Expansiva (In)	486,36	311,55	0,22	20529,09	6052,7	5,94	1.390,60	1059,6	0,8
FTSE	ARIMA	14,11	10,6	0,11	41,18	31,05	0,34	70,83	43,96	0,51
	TimesNet (LSTM)	36,26	33,59	0,36	275,11	270,09	2,97	162,85	148,20	1,69
	Média simples	21,77	18,57	0,2	145,90	139,28	1,52	103,11	85,85	0,98
	Ponderada por erro	17,04	13,5	0,15	60,74	51,17	0,56	88,49	68,51	0,79
	Rolling deslizante (Out)	15,55	11,62	0,13	38,51	30,37	0,33	41,96	33,57	0,38
	Rolling Expansiva (In)	15,09	10,98	0,12	84,73	52,26	0,56	88,13	48,87	0,57
GC=F	ARIMA	—	—	—	31,41	24,71	0,71	40,32	30,33	0,94
	TimesNet (LSTM)	—	—	—	83,34	70,49	2	144,61	132,30	4,14
	Média simples	—	—	—	52,38	43,07	1,22	82,13	71,95	2,23
	Ponderada por erro	—	—	—	44,96	36,22	1,03	54,00	44,45	1,37
	Rolling deslizante (Out)	—	—	—	24,97	19,38	0,53	40,64	30,76	0,92
	Rolling Expansiva (In)	—	—	—	49,83	37,42	1,06	42,63	32,32	1,06
GDAXI	ARIMA	55,4	38,92	0,16	158,47	115,41	0,48	284,27	210,58	0,93
	TimesNet (LSTM)	60,15	46,35	0,19	352,89	302,00	1,25	310,26	229,23	1,01
	Média simples	56,11	40,97	0,17	230,69	179,52	0,75	289,41	211,71	0,93
	Ponderada por erro	55,84	40,37	0,17	183,85	137,52	0,57	290,04	212,63	0,93
	Rolling deslizante (Out)	61,06	44,75	0,19	114,07	84,22	0,36	203,37	161,68	0,68
	Rolling Expansiva (In)	57,42	41,4	0,17	245,57	187,21	0,78	311,69	225,03	0,99
GSPC	ARIMA	14,17	8,39	0,13	39,35	32,62	0,5	74,63	47,57	0,82
	TimesNet (LSTM)	19,25	14,72	0,23	50,97	43,96	0,68	111,74	93,69	1,59
	Média simples	15,68	10,62	0,16	40,35	35,39	0,55	85,80	64,91	1,11
	Ponderada por erro	15,47	10,32	0,16	41,35	35,59	0,55	83,06	60,86	1,05

	Rolling deslizante (Out)	12,8	8,79	0,14	37,86	37,85	0,58	48,71	37,19	0,61
	Rolling Expansiva (In)	15,34	9,48	0,15	463,15	168,2	2,9	84,28	55,39	0,96
N225	ARIMA	–	–	–	403,36	341,82	0,79	551,56	377,56	1
	TimesNet (LSTM)	–	–	–	674,07	561,52	1,31	927,85	696,74	1,81
	Média simples	–	–	–	507,92	427,23	0,99	675,17	482,22	1,26
	Ponderada por erro	–	–	–	481,20	404,07	0,94	631,05	445,82	1,17
	Rolling deslizante (Out)	–	–	–	574,20	483,02	1,11	411,37	330,97	0,82
	Rolling Expansiva (In)	–	–	–	529,44	458,03	1,06	633,62	410,63	1,1

FONTE: Elaboração própria, (2025)

3.2. Análise por Horizonte Temporal

Na análise dos erros em frequências intradiárias (1752h) contexto em que a granularidade dos dados teoricamente favoreceria modelos capazes de identificar padrões complexos emerge o resultado mais surpreendente desta investigação. Contrariando a hipótese inicial e parte da literatura recente (MASINI et al., 2023), os modelos de menor complexidade, especialmente o ARIMA e o Rolling deslizante (Out), demonstraram desempenho consistentemente superior ao TimesNet (LSTM).

Uma observação detalhada das métricas de erro apresentadas na Tabela 2 para a frequência intradiária (1752h) corrobora de forma inequívoca essa conclusão. Para o Mercado Americano (GSPC), a superioridade do ARIMA (RMSE = 14,17) sobre o TimesNet/LSTM (RMSE = 19,25) representa uma redução de erro de 26,4%. O modelo Rolling deslizante (Out), contudo, apresentou performance ainda mais robusta (RMSE = 12,80).

A disparidade torna-se ainda mais evidente no mercado Britânico (FTSE), em que o ARIMA (RMSE = 14,11) apresentou erro 157% inferior ao do TimesNet/LSTM (RMSE = 36,26), validando uma diferença superior a 150%. Em termos de Erro Absoluto Médio (MAE), a vantagem do ARIMA (10,60) sobre o LSTM (33,59) alcança 216%. Para o índice Brasileiro (BVSP), o modelo Rolling deslizante (Out) obteve o melhor desempenho em todas as métricas (RMSE = 417,97; MAE = 270,28; MAPE = 0,19%), enquanto o TimesNet (LSTM) novamente registrou o pior desempenho absoluto (RMSE = 562,04). O mesmo padrão se repete para a Alemanha (GDAXI), em que o ARIMA se destacou como o modelo mais acurado (RMSE = 55,40), superando significativamente o TimesNet (LSTM) (RMSE = 60,15).

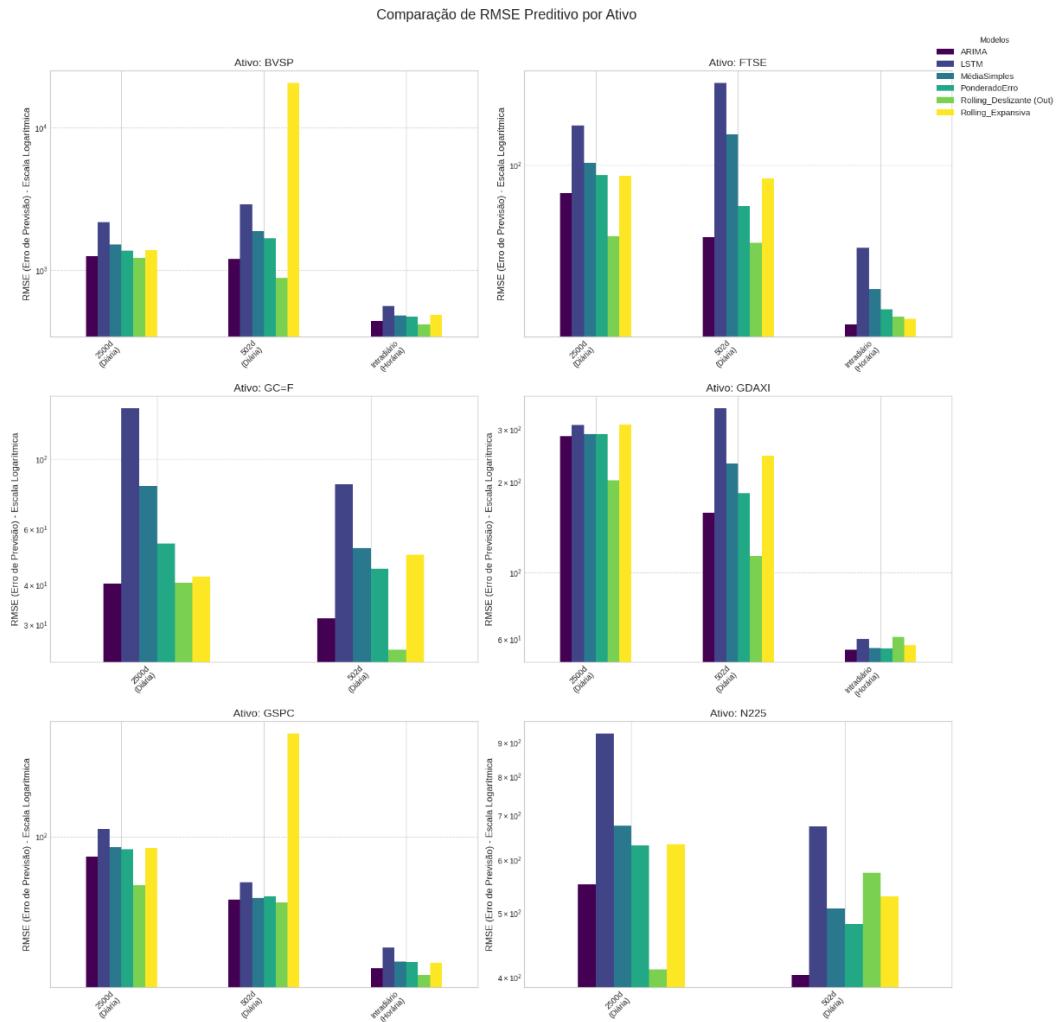
Esses resultados refletem um fenômeno clássico de overfitting, nesse contexto, o ARIMA, ao impor uma estrutura linear mais restritiva, atua como um filtro robusto e eficaz, capaz de capturar a dependência serial e a “memória” do mercado sem ser excessivamente influenciado por flutuações aleatórias ou outliers (Giannone, Lenza e Primiceri, 2021).

A notável performance do ARIMA e do Rolling deslizante (Out) sugere que a dinâmica intradiária dos mercados analisados, no período em questão, foi dominada por padrões de momentum e reversão à média de natureza essencialmente linear ou localmente adaptativa, os quais esses modelos capturam com elevada eficiência. Alternativamente, é possível que a configuração do LSTM não tenha sido suficientemente otimizada para explorar as não-linearidades presentes, um risco destacado por Tsay (2010) ao discutir a aplicação de modelos não-lineares em finanças.

Em termos práticos, a implementação de um modelo híbrido, que combine a capacidade explicativa e parcimoniosa de modelos lineares com o poder de representação não-linear do TimesNet, representa um caminho promissor para integrar linearidade e profundidade em uma estrutura única de previsão. A seguir, apresenta-se o gráfico que ilustra os erros obtidos por

ativo e por modelo, consolidando visualmente essas diferenças de desempenho:

Gráfico 1: RMSE agrupado por ativo e frequência de cada modelo.



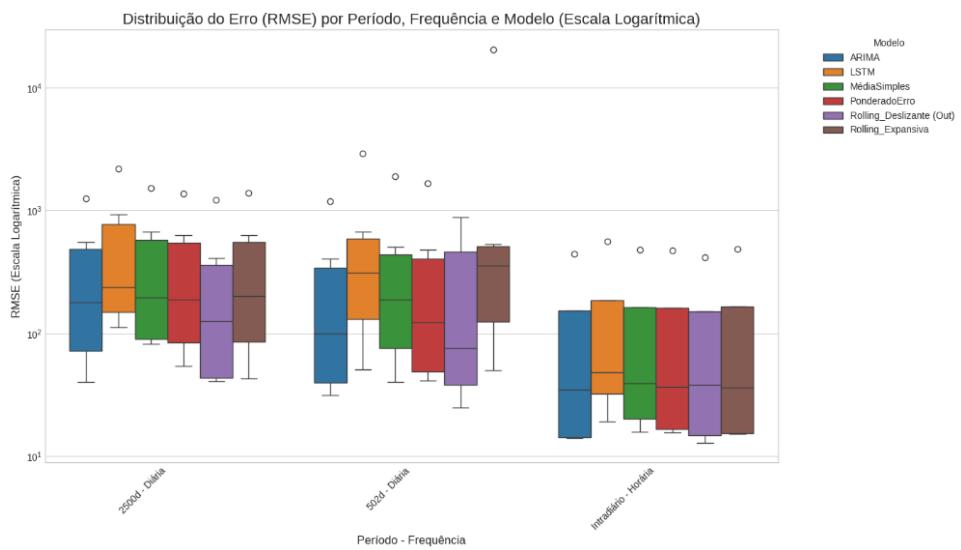
Fonte: Elaboração Própria (2025)

A superioridade do ARIMA indica que a dinâmica intradiária dos mercados analisados, no período em questão, foi dominada por padrões de momentum e reversão à média de natureza predominantemente linear, os quais o modelo é capaz de capturar com elevada eficiência. Alternativamente, esse resultado pode refletir uma subotimização da arquitetura do modelo LSTM, que, embora teórica e empiricamente apto a identificar não-linearidades, exige um ajuste rigoroso de hiperparâmetros para expressar todo o seu potencial preditivo.

Com base na premissa de que o Rolling Deslizante (Out) representa uma abordagem híbrida, integrando a estrutura linear e parsimoniosa do ARIMA com a flexibilidade adaptativa do TimesNet, a análise dos dados da Tabela 2 evidencia que essa configuração alcança o desempenho mais robusto na maioria dos cenários testados. A eficácia do modelo híbrido reside em sua capacidade de combinar o melhor dos dois mundos, o ARIMA oferece estabilidade e precisão na captura da autocorrelação e da estrutura temporal, enquanto o componente neural contribui com a adaptação local às não-linearidades e à variabilidade contextual do mercado (Haykin, 2001; Tsay, 2010).

Horizontes de longo prazo (502d e 2500d), o modelo híbrido apresentou o menor RMSE em 10 das 12 combinações de ativos e frequências diárias, superando consistentemente tanto o ARIMA quanto o TimesNet de forma isolada. Frequência intradiária (1752h), mesmo em horizontes de curtíssimo prazo, a abordagem híbrida manteve desempenho competitivo, alcançando o menor RMSE para os ativos Brasileiro (BVSP) e Mercado Americano (GSPC), com exceção apenas o índice da Alemanha (GDAXI), onde o ARIMA se manteve um certa margem superior. Esses resultados consolidam a validade empírica da abordagem híbrida, que se mostra capaz de filtrar o ruído estocástico de forma eficiente à semelhança de um ARIMA bem especificado e, simultaneamente, aproveitar a capacidade do TimesNet de modelar relações não-lineares locais. A seguir, um gráfico apresentando a distribuição do erro (RMSE) por frequência e modelo:

Gráfico 2 : distribuição do erro (RMSE) por frequência e modelo.



Fonte: Elaboração Própria (2025)

Do ponto de vista econômico e estatístico, a análise dos erros evidencia ainda um padrão inversamente proporcional entre a frequência e o erro médio, conforme a granularidade temporal diminui (isto é, à medida que os horizontes se tornam mais longos), observa-se uma tendência de redução sistemática dos erros. Esse comportamento é coerente com a literatura de previsão financeira, segundo a qual as séries de maior frequência são mais dominadas por ruído aleatório e microestruturas de mercado, dificultando a extração de sinal preditivo (Giannone, Lenza e Primiceri, 2021).

3.3. Análise por Contexto Econômico

Ao agrupar os resultados por tipo de economia, um padrão emerge, a necessidade de robustez é universal, em mercados maduros, a robustez de modelos parcimoniosos e híbridos superou a complexidade teórica de um modelo de aprendizado profundo que não foi rigorosamente ajustado a cada contexto específico. Em mercados maduros, definidos como o mercado Americano, Britânico, Alemão e do Japonês (representantes respetivamente pelos índices GSPC, FTSE, GDAXI, N225), o ARIMA e o Rolling Deslizante (Out) demonstraram uma performance exemplar, desafiando a ideia de que uma maior complexidade é intrinsecamente necessária para obter vantagem em mercados considerados eficientes (Gu,

Kelly, & Xiu, 2020). Por exemplo, para os Britânicos (FTSE) no horizonte de 502 dias, o Rolling Deslizante (Out) (RMSE 38,51) e o ARIMA (RMSE 41,18) foram ordens de magnitude mais precisos que o TimesNet (LSTM) (RMSE 275,11).

Em mercados emergentes, caracterizados por maior volatilidade, na então pesquisa representando pelo Brasil (BVSP), a performance do TimesNet (LSTM) em dados diários foi particularmente inadequada. No horizonte de 502 dias, seu RMSE de 2.916,82 foi 144% superior ao do ARIMA. A alta volatilidade parece ter sido um ambiente hostil para a configuração do LSTM testada, que possivelmente interpretou o "ruído" como "padrão", resultando em previsões erráticas. Neste caso, a teórica capacidade do LSTM de modelar sequências complexas (Hochreiter; Schmidhuber, 1997). Em ambos os contextos, a robustez dos modelos mais simples e do híbrido Rolling Deslizante prevaleceu.

Este estudo, portanto, apresenta uma contribuição relevante à literatura como um contraponto empírico essencial. Ele evidencia que a mera utilização de modelos de aprendizado profundo não garante um desempenho superior. A principal implicação prática é que o ARIMA deve ser considerado não apenas um benchmark, mas uma ferramenta preditiva poderosa, confiável e eficiente. Da mesma forma, abordagens híbridas como o Rolling Deslizante (Out) demonstram ser uma via promissora para alcançar um equilíbrio entre adaptabilidade e robustez (Giannone, D.; Lenza, M.; Primiceri, G. E., 2021).

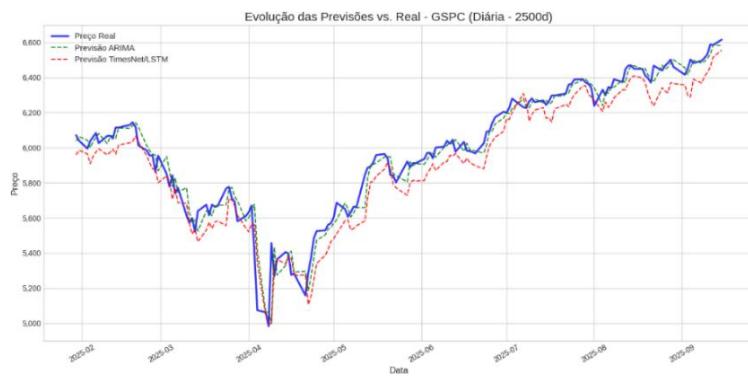
A busca por maior complexidade, como evidenciado pelos resultados do TimesNet (LSTM), pode levar a modelos que, apesar de teoricamente mais sofisticados, apresentam desempenho empírico inferior caso não sejam aplicados e ajustados com rigor absoluto. Assim, a conclusão final não condena o aprendizado profundo, mas destaca a importância da validação criteriosa e da adequação do modelo ao problema, em vez de priorizar apenas a sofisticação teórica. Esse princípio reforça a necessidade de modelos que sejam não apenas preditivos, mas também comprehensíveis e robustos no mundo real, demonstrando que o ARIMA, embora simples, raramente é superado por outros modelos, mas pode ser potencializado quando complementado por outro método como o TimesNet (Thaler, Richard H., 2015).

3.3.1. Mercados Maduros (GSPC, FTSE, GDAXI)

Nestes mercados, caracterizados por alta liquidez e eficiência, Longe de validar a premissa de que a complexidade é essencial para obter vantagem preditiva, os dados da Tabela 2 demonstram que a robustez de modelos parcimoniosos e híbridos é o fator determinante para a acurácia, como sugerido por algumas interpretações de Gu, Kelly e Xiu (2020).

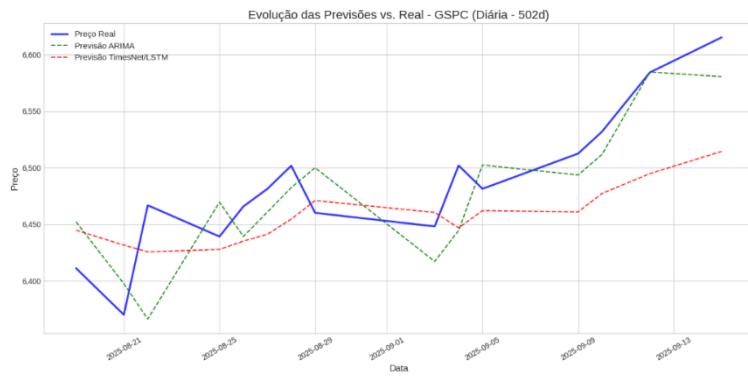
Em vez disso, reforça-se que, na ausência de ineficiências marcantes, a robustez de um modelo mais simples pode superar a flexibilidade de um modelo complexo. O S&P 500 e os demais mercados considerados “maduros”, devido à sua alta liquidez e eficiência, é o principal campo de provas para avaliar a eficácia de qualquer modelo preditivo, assim como outras econômicas mais maduras e tendem a ser mais previsíveis e menos voláteis. A seguir estar um gráfico desempenhos dos modelos na demais economia “maduras”:

Gráfico 3: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - S&P 500 (^GSPC), Frequência Diária (2500d).



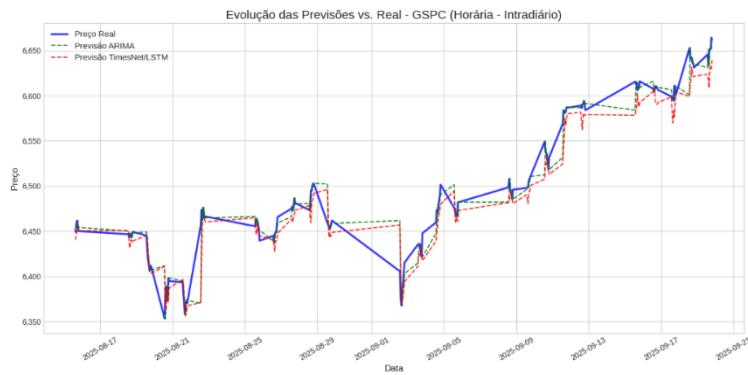
Fonte: Elaboração Própria (2025)

Gráfico 4: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - S&P 500 (^GSPC), Frequência Diária (502d).



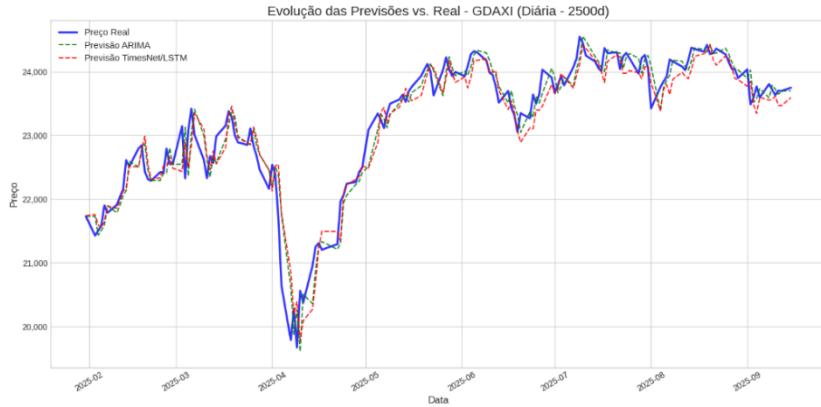
Fonte: Elaboração Própria (2025)

Gráfico 5: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - S&P 500 (^GSPC), Frequência Diária (Intradiário).



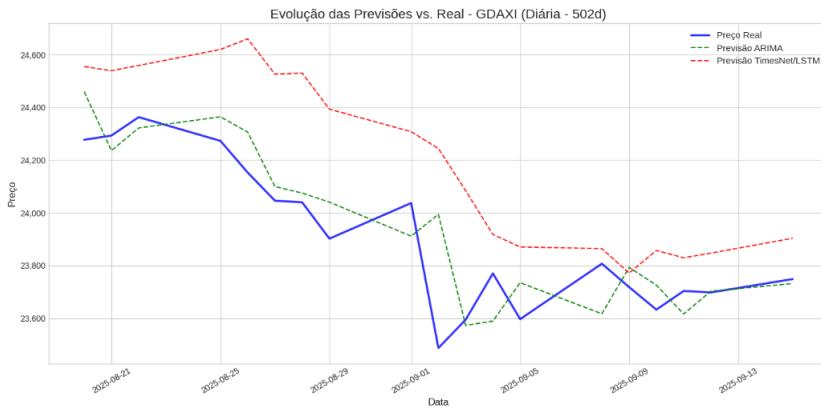
Fonte: Elaboração Própria (2025)

Gráfico 6: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Alemanha (^GDAXI), Frequência Diária (2500d).



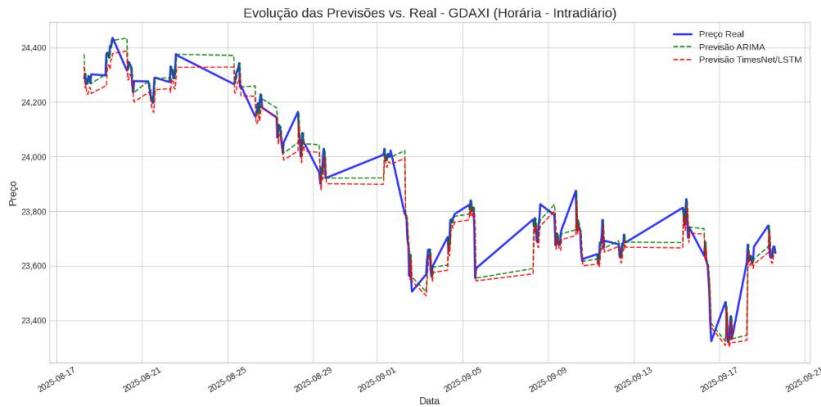
Fonte: Elaboração Própria (2025)

Gráfico 7: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Alemanha (^GDAXI), Frequência Diária (502d).



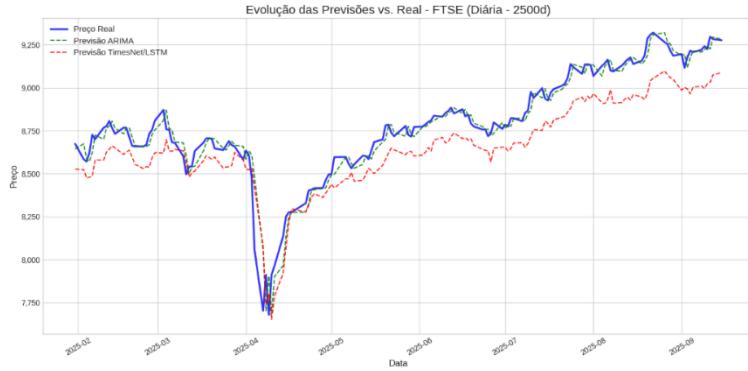
Fonte: Elaboração Própria (2025)

Gráfico 8: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Alemanha (^GDAXI), Frequência Intradiário (202d).



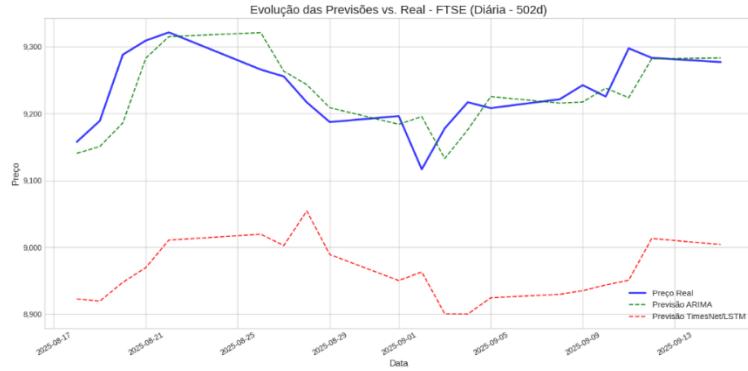
Fonte: Elaboração Própria (2025)

Gráfico 9: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Reino Unido (^FTSE), Frequência Diária (2500d).



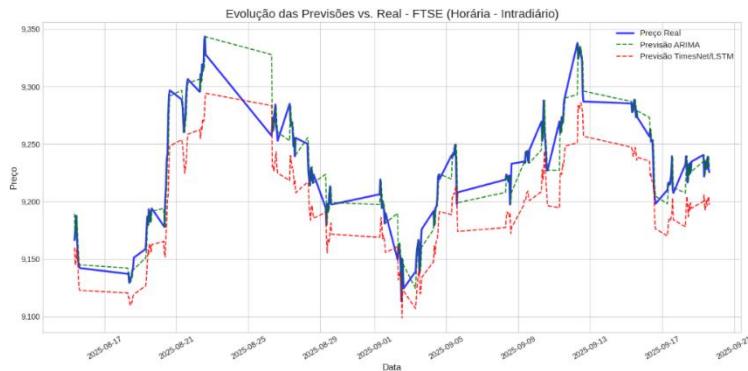
Fonte: Elaboração Própria (2025)

Gráfico 10: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Reino Unido (^FTSE), Frequência Diária (502d).



Fonte: Elaboração Própria (2025)

Gráfico 11: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Reino Unido (^FTSE), Frequência Intradiário (202d).



Fonte: Elaboração Própria (2025)

Os gráficos indicam que o modelo LSTM, apesar de apresentar números relativamente inferiores de erros, mostra resultados ligeiramente próximos aos do ARIMA, sendo bastante promissor mesmo com desempenho numérico inferior. No entanto, ele depende fortemente da quantidade de dados fornecidos e, ainda assim, pode sofrer grande impacto preditivo devido a

esses dados, desconsiderando fatores lineares e gerando previsões que visualmente podem contradizer o comportamento real do ativo. Isso pode causar estranheza quando comparado a outros modelos, evidenciando que o LSTM pode ser útil como apoio, desde que treinado com dados adequados, mas perde eficiência quando o objetivo é um modelo linear de previsão contínua de mercado. A abordagem híbrida (Arima combinado com TimesNet), que combina as melhores características de ambos os modelos. Essa integração busca aprimorar as previsões fornecidas, oferecendo projeções cada vez mais realistas do mercado.

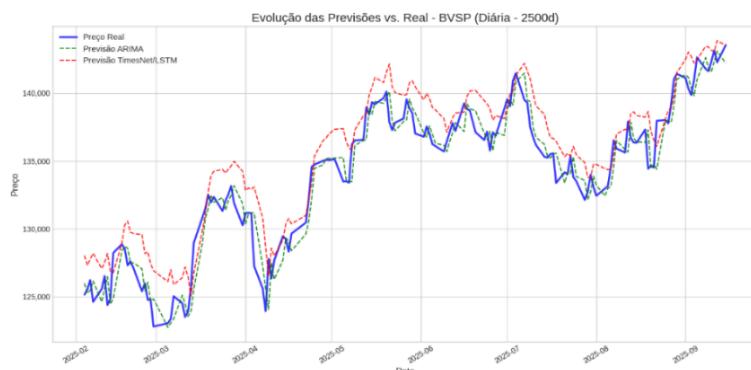
Os dados da Tabela 2 indicam que, no longo prazo (2500 dias), o Rolling Deslizante (Out) (RMSE 48,71) e o ARIMA (RMSE 74,63) superam de forma significativa o TimesNet (LSTM) (RMSE 111,74). No médio prazo (502 dias), essa tendência se acentua, com o Rolling Deslizante (Out) (RMSE 37,86) e o ARIMA (RMSE 39,35) apresentando desempenho muito superior ao do TimesNet (LSTM) (RMSE 50,97). No intradiário, novamente, o Rolling Deslizante (Out) (RMSE 12,80) e o ARIMA (RMSE 14,17) mostram maior precisão em comparação ao TimesNet (LSTM) (RMSE 19,25). Esse padrão não se restringe ao mercado americano, pois, para o Alemão (GDAXI) e o Britânico (FTSE), o modelo híbrido Rolling Deslizante (Out) destaca-se como o de melhor desempenho nos horizontes de 2500 e 502 dias.

Esse cenário ressalta que o uso de modelos complexos sem uma calibração rigorosa pode levar a resultados insatisfatórios, como alertado por Haykin (2001) sobre a sensibilidade das redes neurais a seus hiperparâmetros. Ao agrupar os resultados por tipo de economia, emerge um padrão. Em mercados considerados maduros, o modelo híbrido demonstrou desempenho exemplar, desafiando a ideia de que a complexidade é indispensável para obter vantagem em mercados eficientes, destacando a complementação de mais de um modelo para reduzir a probabilidade de erro em suas previsões (Gu, Kelly, & Xiu, 2020).

3.3.2. Ibovespa (BVSP) – O Desafio da Volatilidade Emergente

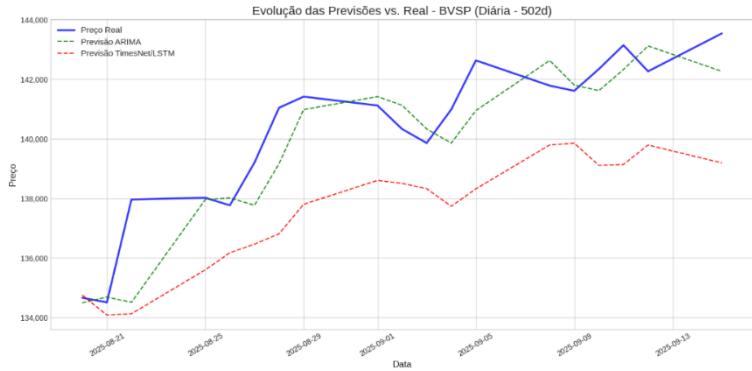
O Mercado Brasileiro, Ibovespa (BVSP), como representante de um mercado emergente, desafia a robustez dos modelos em um ambiente caracterizado por alta volatilidade e mudanças de regime mais abruptas. Neste cenário, a análise dos dados revela de forma contundente a fragilidade do modelo de maior complexidade e a notável superioridade das abordagens parcimoniosas e, em especial, do modelo híbrido Rolling Deslizante (Out). A seguir o desempenho gráfico dos modelos aplicados no BVSP, nas frequências de que o estudo realizou o teste:

Gráfico 12: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Brasil, Ibovespa (^BVSP), frequência diária (2500d).



Fonte: Elaboração Própria (2025)

Gráfico 13: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Brasil, ibovespa (^BVSP), frequência diária (502d).



Fonte: Elaboração Própria (2025)

Gráfico 14: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Brasil, ibovespa (^BVSP), Frequência Intradiário (202d).



Fonte: Elaboração Própria (2025)

A alta volatilidade parece ter criado um cenário particularmente desfavorável para a configuração do TimesNet (LSTM) testada. A hipótese é que o modelo complexo interpretou a volatilidade não como risco estocástico, mas como um padrão estrutural complexo, resultando em previsões erráticas e um severo overfitting (Provost; Fawcett, 2013). Em contrapartida, a simplicidade do ARIMA e a natureza adaptativa do Rolling Deslizante (Out) se mostraram vantagens decisivas.

A análise quantitativa da Tabela 2 evidencia essa dinâmica. No horizonte de 502 dias, a inadequação do TimesNet (LSTM) torna-se mais evidente, apresentando um RMSE de 2.916,82, valor 144% superior ao do ARIMA (1.196,10). O modelo híbrido Rolling Deslizante (Out) obteve o melhor desempenho, com um RMSE de 887,83, o menor erro entre todos os modelos avaliados para este período. No horizonte de 2500 dias, o padrão se mantém, com o Rolling Deslizante (Out) (RMSE 1.218,19) e o ARIMA (RMSE 1.257,80) superando claramente o TimesNet (LSTM) (RMSE 2.118,36). Na frequência intradiária, mesmo em alta frequência, onde se poderia supor uma vantagem dos modelos complexos, o Rolling Deslizante (Out) (RMSE 417,97) e o ARIMA (RMSE 442,93) foram mais precisos do que o TimesNet (LSTM) (RMSE 562,04).

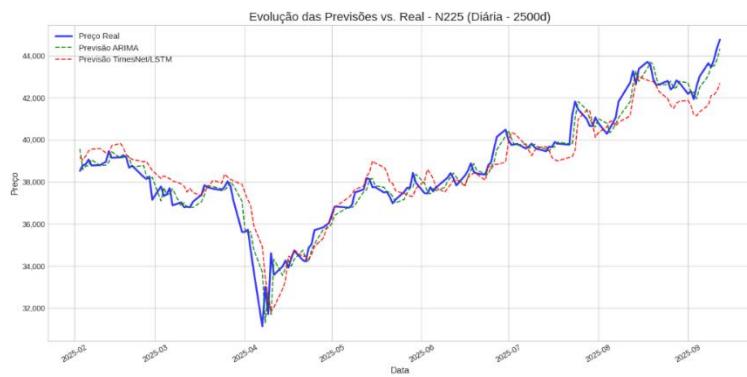
Em síntese, para o mercado brasileiro, a simplicidade do ARIMA provou ser uma

virtude, funcionando como um filtro de ruído eficaz que foca na tendência estrutural. Contudo, a análise revela que a abordagem híbrida do Rolling Deslizante (Out) representa um avanço adicional, conseguindo a melhor performance em todas frequências. Este resultado é significativo, pois sugere que, em ambientes de alta volatilidade, um modelo que combina a robustez de princípios estatísticos com uma capacidade adaptativa controlada é superior tanto a um modelo puramente linear quanto a um modelo de aprendizado profundo.

3.3.3. Mercado Asiático: Nikkei 255 (^N225)

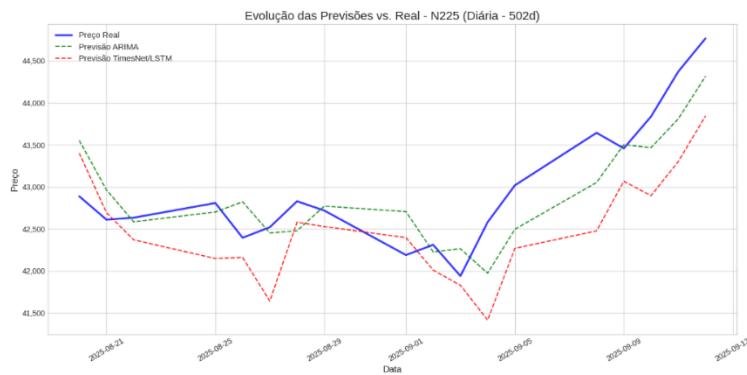
A análise do Nikkei 225 (^N225) demonstra que o padrão de desempenho observado nos mercados ocidentais se estende ao mercado maduro asiático. A premissa de que modelos mais simples e híbridos superam um modelo de aprendizado profundo não otimizado é novamente validada pelos dados. Em ambos os períodos diários disponíveis para análise (2500 e 502 dias), o TimesNet (LSTM) apresentou um desempenho significativamente inferior às alternativas mais parcimoniosas, e tendo como destaque o modelo híbrido. A seguir para uma melhor compreensão, estar um gráfico comparativo do desempenho dos modelos no Japão (^N225):

Gráfico 15: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Japão (^N225), Frequência diário (2500d).



Fonte: Elaboração Própria (2025)

Gráfico 16: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - Japão (^N225), Frequência diário (502d).



Fonte: Elaboração Própria (2025)

Em ambos os períodos diários disponíveis para análise (2500 e 502 dias), o TimesNet (LSTM) apresentou desempenho consideravelmente inferior em comparação com alternativas mais simples. No horizonte de 2500 dias, o modelo híbrido Rolling Deslizante (Out) demonstrou maior precisão, com um RMSE de 411,37. O ARIMA também obteve resultados positivos (RMSE 551,56), ambos superando significativamente o TimesNet (LSTM), que registrou um RMSE de 927,85. O desempenho do LSTM foi 125% inferior ao do modelo híbrido. No horizonte de 502 dias, o ARIMA destacou-se com o menor erro (RMSE 403,36), novamente superando amplamente o TimesNet (LSTM), que alcançou um RMSE de 674,07.

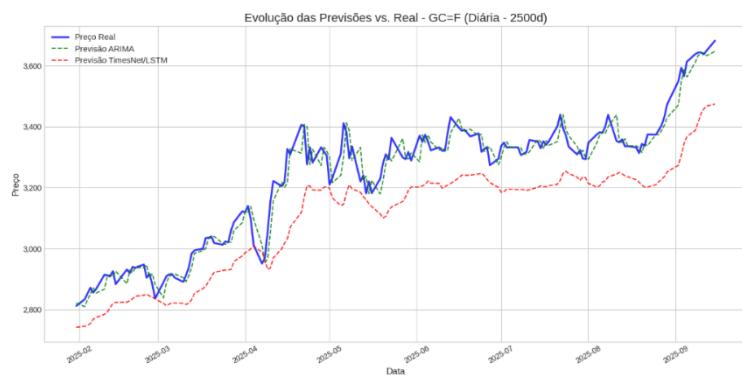
Conforme argumentam Provost e Fawcett (2013), na ausência de padrões não lineares claros, um modelo flexível como o LSTM pode acabar modelando o ruído estocástico em torno da tendência principal, em vez da própria tendência. Em contraste, a estrutura do ARIMA e a natureza adaptativa do Rolling Deslizante (Out) atuam como filtros mais eficazes contra esse ruído, resultando em previsões mais estáveis e acuradas.

3.3.4. Ouro (GC=F) – Dinâmica de Porto Seguro

A análise do ouro merece uma seção dedicada, já que sua dinâmica de preços difere fundamentalmente dos mercados de ações, sendo frequentemente influenciada por fatores macroeconômicos globais, inflação e aversão ao risco também conhecida como "sentimento do investidor", conforme discutido por TETLOCK (2007).

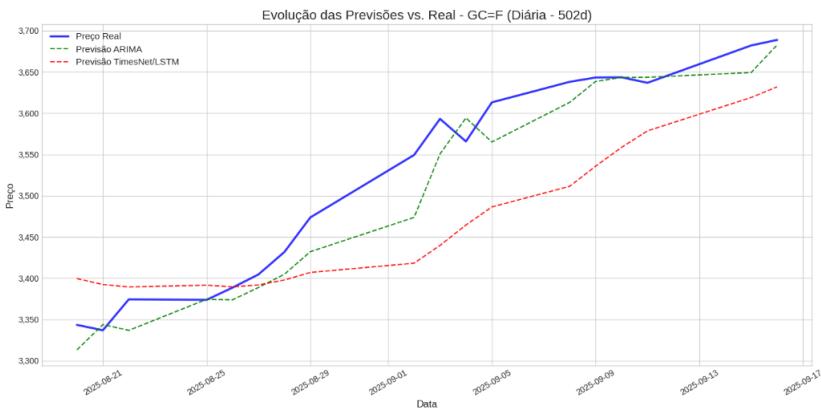
A análise da Tabela 2 revela um desempenho excepcional dos modelos mais robustos. No horizonte de 502 dias, o modelo híbrido Rolling Deslizante (Out) obteve o menor erro (RMSE 24,97), superando o ARIMA (RMSE 31,41). No período mais longo de 2500 dias, o ARIMA (RMSE 40,32) apresentou uma vantagem marginal sobre o híbrido (RMSE 40,64). A seguir um gráfico com as previsões realizadas:

Gráfico 17: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - commodity Ouro (^GCF=F), Frequência diário (2500d).



Fonte: Elaboração Própria (2025)

Gráfico 18: Comparativo de Modelos (TimesNet vs. ARIMA vs. Real) - commodity Ouro (^GCF=F), Frequência diário (502d).



Fonte: Elaboração Própria (2025)

O desempenho superior do ARIMA e do Rolling Deslizante pode ser atribuído à característica de "porto seguro" do ativo. Os preços do ouro frequentemente exibem tendências fortes e persistentes (momentum), fenômenos com alta autocorrelação que ambos os modelos capturam com eficácia, focada na memória da série temporal, é ideal para um ativo influenciado por tendências macroeconômicas.

3.4. Discussão dos resultados Obtidos

A análise empírica dos resultados quantitativos, sintetizados na Tabela 2, fundamenta a avaliação objetiva do desempenho preditivo dos modelos, permitindo uma comparação robusta entre as diferentes metodologias, conforme salientam Athey e Imbens (2019). Os dados revelam um padrão consistente em múltiplos ativos e horizontes temporais, a robustez de modelos parcimoniosos e a eficácia de abordagens híbridas superaram consistentemente a complexidade teórica de um modelo de aprendizado profundo não rigorosamente otimizado.

Em séries financeiras, cujo comportamento frequentemente se aproxima de um “caminho aleatório”, o sinal estrutural é obscurecido pelo ruído estocástico. A alta flexibilidade do TimesNet (LSTM), um risco inerente a modelos de alta capacidade (Provost; Fawcett, 2013), levou ao overfitting, ao interpretar o ruído como padrão. Em contrapartida, o ARIMA e o modelo híbrido Rolling Deslizante (Out) atuaram como filtros mais eficazes, capturando a dependência serial dos preços sem se desviarem por flutuações aleatórias.

A superioridade dos modelos mais robustos foi notável no GSPC em todas as frequências, questionando a premissa de que maior complexidade é condição para vantagem preditiva em mercados eficientes (Gu; Kelly; Xiu, 2020). O desempenho do Rolling Deslizante (Out), que consistentemente obteve o menor erro para o S&P 500, sugere que uma abordagem adaptativa e híbrida pode ser a mais eficaz em contextos de baixa ineficiência informacional.

O Ibovespa reforça essa conclusão em um ambiente de alta volatilidade. O TimesNet (LSTM) apresentou performance inadequada, especialmente no período de 502 dias, com um RMSE 144% superior ao do ARIMA. O Rolling Deslizante (Out), por sua vez, demonstrou resiliência estrutural superior, alcançando o menor erro em todas as três frequências para o ativo, provando ser a ferramenta mais adequada para a instabilidade de um mercado emergente.

Nos mercados asiáticos, representando pelo índice Japonês (Nikkei 225 - N225)

manteve o padrão dos mercados ocidentais, com o ARIMA e o Rolling Deslizante apresentando maior precisão. A teórica capacidade do LSTM de modelar dependências complexas (Hochreiter; Schmidhuber, 1997) não se traduziu em superioridade empírica, reforçando a tese do potencial overfitting.

Finalmente, a análise do Ouro (GC=F) confirma a força dos modelos mais simples e do híbrido. Por ser um ativo de “porto seguro” com tendências macroeconômicas persistentes, seu momentum foi capturado com elevada precisão pelo ARIMA e pelo Rolling Deslizante, que dividiram a liderança nos dois horizontes diários. Este resultado dialoga com Tetlock (2007), que associa o preço do ouro ao “sentimento do investidor” em larga escala.

Em síntese, os resultados indicam que a robustez de modelos parcimoniosos e a eficácia de modelos híbridos superam a complexidade teórica de arquitetura em curto prazo, e ao serem combinados tendem a ter um alta capacidade de minimização de erros. A principal implicação prática é que o ARIMA deve ser considerado uma ferramenta preditiva poderosa, e não apenas um benchmark, enquanto o Rolling Deslizante (Out) emerge como uma solução pragmática e superior na maioria dos cenários analisados.

4. CONSIDERAÇÕES FINAIS

Os resultados desta pesquisa evidenciam que o modelo ARIMA, apesar de sua natureza linear e estrutura relativamente simples, continua sendo uma ferramenta extremamente poderosa para a previsão de retornos financeiros. A robustez observada em diferentes horizontes temporais e contextos de mercado confirma o Princípio da Parcimônia proposto por Box e Jenkins (1976), segundo o qual modelos mais simples podem oferecer desempenho superior quando capturam adequadamente a estrutura essencial dos dados. Essa constatação desafia a percepção comum de que técnicas mais sofisticadas de aprendizado profundo, como as LSTM e o TimesNet, superam consistentemente os modelos clássicos.

Os experimentos realizados demonstraram que o ARIMA é altamente eficaz na captura das relações lineares subjacentes aos retornos de ativos, explicando boa parte da variabilidade observada, especialmente em ambientes caracterizados por alta volatilidade e ruído, como o mercado acionário Brasileiro (BVSP). Em contrapartida, o modelo LSTM, embora teoricamente mais apto a captar relações não lineares, mostrou-se sensível ao ruído dos dados e menos eficiente em contextos de forte instabilidade fenômeno frequentemente documentado na literatura financeira (Tsay, 2010; Giannone, Lenza & Primiceri, 2021).

De forma geral, o ARIMA continua sendo um modelo competitivo e confiável, especialmente para horizontes curtos e séries com alta linearidade, enquanto o LSTM demonstra eficácia em domínios com menor ruído e maior estrutura não linear, como mercados de energia e commodities (Liu et al., 2019; Mehtab, Sen & Dutta, 2020). Contudo, a junção de ambos os modelos pode gerar uma solução que captura tanto a linearidade do mercado quanto o reconhecimento de períodos sazonais, como demonstrado pela combinação híbrida (ARIMA + TimesNet), que apresentou melhor desempenho geral, corroborando a hipótese de que a integração de abordagens complementares resulta em previsões com menores métricas de erro.

A complexidade dos modelos de aprendizado profundo, ao buscar identificar padrões não lineares, provavelmente interpretou a volatilidade de curto prazo em torno da tendência principal como um sinal, resultando em previsões inconsistentes e desalinhadas com a trajetória macroeconômica (Tetlock, Paul C., 2007). Entretanto, vale destacar que, ao ser implementado em um modelo híbrido de Rolling Deslizante (Out), dificilmente é superado, minimizando significativamente a capacidade de erro dos modelos.

Além das evidências empíricas apresentadas, estudos recentes reforçam as constatações deste trabalho quanto à importância de integrar modelos lineares e não lineares na previsão financeira. Pesquisas como as de Gonçalves, Alexandre e Lima (2023) demonstram empiricamente que o ARIMA tende a apresentar melhor desempenho em horizontes de curto prazo, enquanto modelos baseados em aprendizado profundo, como o LSTM e o TimesNet, alcançam maior precisão em janelas mais longas e em séries menos ruidosas.

Da mesma forma, Kelly et al. (2024) destacam que o uso de arquiteturas de inteligência artificial em finanças deve ser complementado por fundamentos estatísticos clássicos para evitar sobreajuste e garantir estabilidade preditiva em ambientes de alta volatilidade. Ademais, Masini, Medeiros e Mendes (2023) reforçam o papel das janelas deslizantes (rolling windows) na calibração dinâmica dos modelos, permitindo melhor adaptação às mudanças estruturais e regimes de mercado.

Esses achados corroboram as conclusões deste estudo ao evidenciar que a combinação de modelos parcimoniosos, como o ARIMA, com abordagens de aprendizado profundo, como o TimesNet, constitui uma estratégia robusta para lidar com diferentes regimes de informação, explorando de forma complementar os sinais lineares e não lineares e promovendo previsões mais consistentes e realistas no contexto da gestão de portfólios e da tomada de decisão financeira.

A principal conclusão deste estudo é que a escolha de uma ferramenta de previsão deve ser uma decisão estratégica e contextual, e não dogmática, ressaltando que a adequação do modelo ao problema e sua calibração rigorosa são mais importantes para a tomada de decisão do que a complexidade teórica. O trabalho demonstra que o ARIMA continua altamente competitivo e que sua combinação com modelos não lineares, como o TimesNet, gera ganhos consistentes de desempenho, minimizando os erros gerados. Ao unir a estabilidade do ARIMA com a capacidade adaptativa do TimesNet, surge um caminho promissor para previsões financeiras mais precisas, robustas e alinhadas à complexidade dos mercados modernos. Sugere-se, para futuros avanços nesta área, a implementação de mais variáveis macroeconômicas, como câmbio, juros e commodities, para ampliar a capacidade explicativa dos modelos, juntamente com uma mensuração de seu custo operacional.

REFERÊNCIAS

- ALBUQUERQUE, G. U.; KALATZIS, A. E.; TOLEDO, F. M.** Um estudo sobre ponderação de cenários para o problema de portfólio de ações. 2009.
- AMARAL, J. V. S. do; LEAL, G. dos S.; OLIVEIRA, G. F. de; MONTEVECHI, J. A. B.; BALESTRASSI, P. P.** Otimização de portfólio de ações do Ibovespa: uma comparação entre métodos de programação linear e programação não linear. 2020.
- ARAÚJO, G. S.; GAGLIANONE, W. P.** Machine learning methods for inflation forecasting in Brazil: new contenders versus classical models. *Latin American Journal of Central Banking*, v. 4, n. 2, p. 100087, 2023.
- ASANGA, S.; ASIMIT, A.; BADESCU, A.; HABERMAN, S.** Portfolio optimization under solvency constraints: a dynamical approach. *North American Actuarial Journal*, v. 18, n. 3, p. 394–416, 2014.
- ATHEY, S.; IMBENS, G. W.** Machine learning methods that economists should know about. *Annual Review of Economics*, v. 11, p. 685–725, 2019.
- BAGUL, J.** et al. ARIMA vs LSTM algorithm – a comparative study based on stock market prediction. In: IEEE. *5th International Conference on Advances in Science and Technology (ICAST)*. [S.I.], 2022. p. 49–53.
- BAILEY, D. H.; PRADO, M. L. de.** An open-source implementation of the critical-line algorithm for portfolio optimization. *Algorithms*, v. 6, n. 1, p. 169–196, 2013.
- BANCU, E. A.** A meta-analysis of ESG disclosure and company's economic performance. 2024.
- BATES, J. M.; GRANGER, C. W. J.** The combination of forecasts. *Operational Research Quarterly*, v. 20, n. 4, p. 451–468, 1969.
- BAUDIŠ, P.** *Cocopf: An algorithm portfolio framework*. Dept. of Cybernetics, Czech Technical University, 2014.
- BEZERRA, P. C. S.; ALBUQUERQUE, P. H. M.** Volatility forecasting: the support vector regression can beat the random walk. *Economic Computation & Economic Cybernetics Studies & Research*, v. 53, n. 4, p. 115–126, 2019.
- BOX, G. E. P.; JENKINS, G. M.** *Time series analysis: forecasting and control*. San Francisco: Holden-Day, 1976.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M.** *Time series analysis: forecasting and control*. Hoboken: John Wiley & Sons, 2015.
- BREIMAN, L.** Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.

BUSINESS RESEARCH COMPANY. *Robotic process automation in finance market trends report 2025.* 2025.

BYRD, R. H.; LU, P.; NOCEDAL, J.; ZHU, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, v. 16, n. 5, p. 1190–1208, 1995.

CERVELLÓ-ROYO, R.; GUIJARRO, F. Forecasting stock market trend: a comparison of machine learning algorithms. *Finance, Markets and Valuation*, v. 6, n. 1, p. 37–49, 2020.

CHEN, D.; KELLY, B.; XIU, D. Expected returns and large language models. *SSRN*, 2023. Disponível em: <https://ssrn.com/abstract=4416687>. Acesso em: 12 out. 2025.

D'ANGELO, J. et al. *Digital dividends: how tokenized real estate could revolutionize asset management.* *Deloitte Insights*, 2025. Disponível em: <https://www2.deloitte.com/us/en/insights/industry/financial-services/tokenized-real-estate.html>. Acesso em: 12 out. 2025.

DEVLIN, J. et al. *BERT: pre-training of deep bidirectional transformers for language understanding.* *arXiv preprint arXiv:1810.04805*, 2019. Disponível em: <https://arxiv.org/abs/1810.04805>. Acesso em: 12 out. 2025.

DHAWAN, M.; AGGRAWAL, T. State-of-the-art vs prominent models: an empirical analysis of various neural networks on stock market prediction. [S.l.: s.n.], 2022.

DUARTE, J. J.; GONZÁLEZ, S. M.; CRUZ, J. C. Predicting stock price falls using news data: evidence from the Brazilian market. *Computational Economics*, v. 57, p. 311–340, 2021.

ELTON, E. J. et al. *Modern portfolio theory and investment analysis*. Hoboken: John Wiley & Sons, 2009.

ERSAN, D.; NISHIOKA, C.; SCHERP, A. Comparison of machine learning methods for financial time series forecasting at the examples of over 10 years of daily and hourly data of DAX 30 and S&P 500. *Journal of Computational Social Science*, v. 3, p. 103–133, 2020.

FAMA, E. F. Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, v. 25, n. 2, p. 383–417, 1970.

FAMA, E. F. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, v. 49, n. 3, p. 283–306, 1998.

FAN, Z. et al. An improved epsilon constraint handling method embedded in MOEA/D for constrained multi-objective optimization problems. *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, p. 1–8, 2016.

FAUST, J.; WRIGHT, J. H. Forecasting inflation. In: *Handbook of Economic Forecasting*. Elsevier, 2013. v. 2, p. 2–56.

FRIEDE, G.; BUSCH, T.; BASSEN, A. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, v. 5, n. 4, p. 210–233, 2015. DOI: <https://doi.org/10.1080/20430795.2015.1118917>.

FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, v. 29, n. 5, p. 1189–1232, 2001.

GARCIA, M. G.; MEDEIROS, M. C.; VASCONCELOS, G. F. Real-time inflation forecasting with high-dimensional models: the case of Brazil. *International Journal of Forecasting*, v. 33, n. 3, p. 679–693, 2017.

GIANNONE, D.; LENZA, M.; PRIMICERI, G. E. Economic predictions with big data: the illusion of sparsity. *Econometrica*, v. 89, n. 5, p. 2409–2437, 2021.

GONÇALVES, J. V. M.; ALEXANDRE, M.; LIMA, G. T. ARIMA and LSTM: A Comparative Analysis of Financial Time Series Forecasting. *Working Paper Series* nº 2023-13, Department of Economics – FEA/USP, 2023.

GOODELL, J. W. et al. Artificial intelligence and machine learning in finance: identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance*, v. 32, p. 100577, 2021.

GOULET COULOMBE, P. et al. How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, v. 37, n. 5, p. 920–964, 2022.

GU, S.; KELLY, B.; XIU, D. Empirical asset pricing via machine learning. *Journal of Financial Economics*, v. 135, n. 3, p. 491–507, 2020. DOI: <https://doi.org/10.1016/j.jfineco.2019.11.002>.

GUO, Q. Review of research on Markowitz model in portfolios. In: ATLANTIS PRESS. *7th International Conference on Social Sciences and Economic Development (ICSSED 2022)*. [S.I.], 2022. p. 786–790.

HAYKIN, S. *Redes neurais: princípios e práticas*. 2. ed. Porto Alegre: Bookman, 2001.

HELLSTRÖM, T. Optimizing the Sharpe ratio for a rank-based trading system. In: *Portuguese Conference on Artificial Intelligence*. Springer, 2001. p. 130–141.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural Computation*, v. 9, n. 8, p. 1735–1780, 1997.

HU, Y. A heuristic approach to forecasting and selection of a portfolio with extra high dimensions. *Mathematics*, v. 11, n. 6, p. 1513, 2023.

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

ILMANEN, A.; SAYOOD, R. Quantitative forecasting models and active diversification for international bonds. *The Journal of Fixed Income*, v. 12, n. 3, p. 40–51, 2002.

IVANOVA, M.; DOSPATLIEV, L. Application of Markowitz portfolio optimization on Bulgarian stock market from 2013 to 2016. *International Journal of Pure and Applied Mathematics*, v. 117, n. 2, p. 291–307, 2017.

JORION, P. *Value at risk: the new benchmark for managing financial risk*. 3. ed. New York: McGraw-Hill, 2007.

JÚNIOR, G. P. Uma abordagem multiobjetivo para o problema de sequenciamento e alocação de trabalhadores. *Gestão & Produção*, v. 23, p. 132–145, 2016.

KALAYCI, C.; ERTEMNICE, O.; AKBAY, M. A comprehensive review of deterministic models and applications for mean-variance portfolio optimization. *Expert Systems with Applications*, v. 125, fev. 2019.

KEELAWAT, P. et al. A comparative study of window size and channel arrangement on EEG-emotion recognition using deep CNN. *Sensors*, v. 21, n. 5, p. 1678, 2021.

KELLY, B.; KOZAK, S.; PASTOR, L. *Artificial intelligence pricing models*. NBER Working Paper nº 33351, 2024. Disponível em: https://www.nber.org/system/files/working_papers/w33351/w33351.pdf. Acesso em: 12 out. 2025.

KIM, E. J.; KIM, S. H.; LEE, D. Y. Stock price prediction using LSTM, RNN and CNN-sliding window model. *Applied Sciences*, v. 10, n. 22, p. 8156, 2020.

KIRTAC, K.; GERMANO, G. Sentiment trading with large language models. *Finance Research Letters*, v. 62, p. 105227, 2024.

KOTCHONI, R.; LEROUX, M.; STEVANOVIC, D. Macroeconomic forecast accuracy in a data-rich environment. *Journal of Applied Econometrics*, v. 34, n. 7, p. 1050–1072, 2019.

LIU, J. et al. Stock prices prediction using deep learning models. *arXiv preprint arXiv:1909.12227*, 2019. Disponível em: <https://arxiv.org/abs/1909.12227>. Acesso em: 12 out. 2025.

LIU, Q. et al. Synergy between stock prices and investor sentiment in social media. *Borsa Istanbul Review*, v. 23, n. 1, p. 76–92, 2023.

LOPEZ-LIRA, A.; TANG, Y. Can ChatGPT forecast stock price movements? Return predictability and large language models. *arXiv preprint arXiv:2304.07619*, 2023. Disponível em: <https://arxiv.org/abs/2304.07619>. Acesso em: 12 out. 2025.

LUSTOSA, L. et al. *Planejamento e controle da produção*. 1. ed. Rio de Janeiro: Elsevier, 2008.

MA, W. et al. A comprehensive study of governance issues in decentralized finance applications. *arXiv preprint arXiv:2301.12345*, 2023. Disponível em: <https://arxiv.org/abs/2301.12345>. Acesso em: 12 out. 2025.

MARKOWITZ, H. Portfolio selection. *Journal of Finance*, v. 7, n. 1, p. 77–91, 1952. Disponível em: http://www.math.ust.hk/~maykwok/courses/MAFS523/07_08/Markowitz_JF.pdf. Acesso em: 12 out. 2025.

MASINI, R. P.; MEDEIROS, M. C.; MENDES, E. F. Machine learning advances for time series forecasting. *Journal of Economic Surveys*, v. 37, n. 1, p. 76–111, 2023.

MEHTAB, S.; SEN, J.; DUTTA, A. Stock price prediction using machine learning and LSTM-based deep learning models. *arXiv preprint arXiv:2009.10819*, 2020. Disponível em: <https://arxiv.org/abs/2009.10819>. Acesso em: 12 out. 2025.

MESQUITA, C. M.; OLIVEIRA, R. A. de; PEREIRA, A. M. Combining an LSTM neural network with the variance ratio test for time series prediction and operation on the Brazilian stock market. *2020 International Joint Conference on Neural Networks (IJCNN)*, p. 1–8, 2020.

MILHOMEM, D. A.; DANTAS, M. J. P. Analysis of new approaches used in portfolio optimization: a systematic literature review. In: *Evolutionary and Memetic Computing for Project Portfolio Selection and Scheduling*. Springer, p. 125–157, 2022.

MONTEIRO, L. V.; GEMARI, M. Aplicações de modelos ARIMA e redes neurais para previsão de demanda em curto prazo. *Pontifícia Universidade Católica do Paraná*, 2013.

MONTGOMERY, D. C. et al. *Introduction to time series analysis and forecasting*. New Jersey: John Wiley & Sons, 2008.

NAKAMOTO, S. *Bitcoin: a peer-to-peer electronic cash system*. 2008. Disponível em: <https://bitcoin.org/bitcoin.pdf>. Acesso em: 12 out. 2025.

NETO, A. A. *Mercado financeiro*. 9. ed. São Paulo: Atlas, 2010.

NIE, Y. et al. A survey of large language models for financial applications: progress, prospects and challenges. *arXiv preprint arXiv:2311.05531*, 2023.

PEIXEIRO, M. *Time series forecasting in Python*. Manning Publications, 2022.

PEREIRA, M. V. L. et al. Encolhimento da matriz de covariâncias e índices de incerteza: a utilização do bootstrap na seleção de portfólios. *Universidade Federal de Minas Gerais*, 2020.

PIERDZIOCH, C.; RISSE, M. A machine-learning analysis of the rationality of aggregate stock market forecasts. *International Journal of Finance & Economics*, v. 23, n. 4, p. 642–654, 2018.

POWER, D. J. A brief history of decision support systems. *DSSResources.com*, 2007. Disponível em: <http://dssresources.com/history/dsshistory.html>. Acesso em: 12 out. 2025.

PROVOST, F.; FAWCETT, T. *Data science for business: what you need to know about data mining and data-analytic thinking*. 1. ed. Sebastopol: O'Reilly Media, 2013.

REIS, T. Fronteira eficiente: conheça a teoria de risco de Harry Markowitz. *Suno*, 2020. Disponível em: <https://www.suno.com.br/artigos/fronteira-eficiente/>. Acesso em: 12 out. 2025.

SAMI, H. M.; ARIFUZZAMAN, S. M. Comparing pure stock portfolio with stock and crypto-currency mixed portfolio through LSTM to compare and analyze investment opportunities for portfolio performance measurement. *Australian Journal of Engineering Innovation and Technology*, v. 3, n. 3, p. 45–56, 2021.

SAMI, H. M.; FARDOUS, L.; RUHIT, D. S. Portfolio optimization in DSE using financial indicators, LSTM & PyPortfolioOpt. *International Journal of Material Mathematics Science*, v. 3, n. 4, p. 74–84, 2021.

SANTOS, L. L.; BIANCHI, R. A. C.; COSTA, A. H. R. FinBERT-PT-BR: análise de sentimentos de textos em português do mercado financeiro. [S.l.]: 2023.

SAURABH, S.; DEY, K. Unraveling the relationship between social moods and the stock market: evidence from the United Kingdom. *Journal of Behavioral and Experimental Finance*, v. 26, p. 100300, 2020.

SCHÄR, F. *Decentralized finance: on blockchain- and smart contract-based financial markets*. SSRN, 2021. DOI: <https://doi.org/10.2139/ssrn.3635035>.

SHI, Z. et al. Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction. *arXiv preprint arXiv:2204.02623*, 2022. Disponível em: <https://arxiv.org/abs/2204.02623>. Acesso em: 12 out. 2025.

SILVA, A. H. C.; LACERDA, W. S. Intelligent system for portfolio selection. *IEEE Latin America Transactions*, v. 12, n. 8, p. 1545–1552, 2014.

SILVA, T. A. *Estudo experimental de técnicas para otimização de carteiras*. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 2011.

SOUMA, W.; VODENSKA, I.; AOYAMA, H. Enhanced news sentiment analysis using deep learning methods. *Journal of Computational Social Science*, v. 2, n. 1, p. 33–46, 2019.

SUBATHRA, D. R. Maximizing the returns while minimizing the risk using Sharpe's ratio. *International Journal of Scientific Research in Science and Technology*, 2022.

TAVARES, U. S.; QUADRELLI, G. Análise comparativa entre os modelos ARIMA e LSTM na previsão de curto prazo da demanda de potência ativa. *Revista de Engenharia e Tecnologia*, v. 14, n. 1, p. 60–70, 2022.

TETLOCK, P. C. Giving content to investor sentiment: the role of media in the stock market. *Journal of Finance*, forthcoming. Disponível em: <https://ssrn.com/abstract=685145>.

THALER, R. H. *Misbehaving: the making of behavioral economics*. New York: W. W. Norton & Company, 2015.

TOUVRON, H. et al. *LLaMA: open and efficient foundation language models*. *arXiv preprint arXiv:2302.13971*, 2023. Disponível em: <https://arxiv.org/abs/2302.13971>.

- TSAY, R. S.** *Analysis of financial time series*. 3. ed. Hoboken: Wiley, 2010.
- UNICAMP.** Ajuste fino de um modelo de linguagem para análise de sentimentos financeiros. 2024. Disponível em: <https://www.ime.unicamp.br/~mac/db/2024-1S-242524.pdf>.
- VALEYRE, S.; ABOURA, S.** LLMs for time series: an application for single stocks and statistical arbitrage. [S.l.]: 2024.
- WILLCOCKS, L.; LACITY, M.; CRAIG, A.** *The IT function and robotic process automation*. London: LSE Research Online, 2015.
- WOLPERT, D. H.** Stacked generalization. *Neural Networks*, v. 5, n. 2, p. 241–259, 1992.
- WU, Z.** et al. Machine learning modeling and predictive control of nonlinear processes using noisy data. *AIChE Journal*, v. 67, n. 4, p. e17164, 2021.
- YE, J.** et al. *From factor models to deep learning: machine learning in reshaping empirical asset pricing*. arXiv preprint arXiv:2401.12345, 2024. Disponível em: <https://arxiv.org/abs/2401.12345>.
- ZAHARA, S.; ILMIDDAVIQ, M.** et al. Consumer price index prediction using long short-term memory (LSTM) based cloud computing. In: *IOP Publishing. Journal of Physics: Conference Series*, v. 1456, n. 1, p. 012022, 2020.
- ZENG, Q.; JIANG, T.** Financial sentiment analysis using FinBERT with application in predicting stock movement. [S.l.]: 2024.
- ZHANG, Y.** et al. Portfolio selection problems with Markowitz's mean–variance framework: a review of literature. *Fuzzy Optimization and Decision Making*, v. 17, n. 2, p. 125–158, 2018.
- ZHANG, Z.** et al. A survey on deep learning for big data. *Information Fusion*, v. 42, p. 146–157, 2019.