

O USO DE MACHINE LEARNING COMO ALTERNATIVA PARA A PREDIÇÃO DA TAXA DE CÂMBIO

Palavras-Chave: Taxa de Câmbio, Machine Learning, Deep Learning

1. INTRODUÇÃO

Nesta secção são apresentados o contexto — no qual o presente trabalho está inserido — problema de pesquisa e justificativa que delimitam objetivos, hipóteses testadas.

1.1 Contextualização

Na literatura econômica, estabeleceu-se a ideia de que, em um mercado eficiente, todas as informações referentes ao valor de um ativo estão contidas em seu preço (Fama, 1970). Dessa forma, modelos teóricos seriam ineficazes para prever quaisquer movimentos de preço utilizando variáveis empíricas, sendo o valor histórico o melhor preditor do valor futuro. Em outras palavras, o modelo de Random Walk deve apresentar a melhor performance em relação aos demais modelos, quando avaliados em seus termos de erro.

Em seu trabalho, Meese e Rogoff (1983) confirmam essa hipótese ao comparar os modelos teóricos mais robustos da época e verificar que estes foram incapazes de superar o modelo de Random Walk em um horizonte de um a doze meses. Esses resultados ficaram conhecidos na literatura como o puzzle de Meese e Rogoff, estabelecendo o conceito de que nenhum modelo econométrico conseguiria superar o modelo histórico. Em tempos recentes, os avanços envolvendo machine learning possibilitaram uma alternativa à metodologia tradicional de predição de dados e uma possível solução para este obstáculo da literatura.

Os avanços no campo da inteligência artificial permitiram o desenvolvimento de técnicas inovadoras que, ao invés da abordagem clássica de criação de algoritmos pré-definidos executados por computadores, a máquina aprende a reconhecer padrões e tendências através do erro. Essa nova tecnologia, chamada de Machine Learning (ML), tem ampliado os limites na área de previsão de dados, permitindo pesquisadores irem além do método tradicional de definição de variáveis dependentes e independentes. O ML assimila a variação histórica de preços de um ativo e determina seu valor futuro, facilitando a criação de modelos mais robustos para a previsão da taxa de câmbio.

1.2 Justificativa

Devido as características altamente dinâmicas e interdependentes das variáveis macroeconômicas, desenvolver um modelo capaz de prever com exatidão a taxa futura do câmbio representa uma tarefa de complexidade alta. No entanto, ao se utilizar técnicas de machine learning, os modelos podem reconhecer tendências históricas de longos períodos e, a partir disso, gerar resultados precisos.

Como consequência da efetividade destes modelos, formuladores de políticas públicas saberiam de antemão como os resultados afetariam diretamente o câmbio e parte da economia. Pesquisadores poderiam desenvolver teorias e métricas com maior acurácia utilizando como premissa o valor pré-definido de moedas internacionais. Por fim, arbitradores seriam mais eficientes ao negociar ativos em diversos mercados ao redor do mundo, melhorando assim o ajuste dos preços ao seu valor justo.

1.3 Problema de Pesquisa e Objetivo

O principal objetivo deste estudo se concentra em responder a seguinte pergunta: podem modelos de machine learning superar o puzzle de Meese-Rogoff? Para atingir este objetivo, definiu-se as moedas e horizontes de tempos a serem utilizados para a criação dos modelos desejados, e, tendo como base o RMSE as comparou ao modelo de Random Walk. A hipótese estabelecida está alinhada ao problema de pesquisa e a literatura estudada e é definida como: H1: Modelos de machine learning produzem um erro quadrático menor que o modelo de Random Walk.

1.4 Contribuições

Devido à escassez de trabalhos relacionados ao tema, este estudo contribui com a discussão sobre o uso de machine learning como solução para a previsão do câmbio. Introduzindo cinco técnicas recentes e heterogêneas, exploramos o potencial de cada uma delas e sua efetividade no uso envolvendo dados em séries de tempo.

Nossos resultados representam um caminho viável para superar o puzzle de Meese e Rogoff, ainda que os modelos empregados não tenham apresentado performance superior ao modelo de Random Walk, eles demonstram a efetividade de cada modelo em relação ao outro, permitindo que futuros estudos explorem com mais profundidade cada um deles.

2 FUNDAMENTAÇÃO TEÓRICA

Apresenta-se nesta seção, as principais teorias que tem por objetivo suportar as hipóteses desse estudo.

2.1 Hipotese de Mercado Eficiente

Uma das teorias mais importantes no campo econômico-financeiro, é a Hipotese de Mercado Eficiente (EMH) descrita por Fama (1970) em seu trabalho seminal sobre a eficiência dos mercados de capitais. Esse conceito estabelece que em um mercado onde, a qualquer momento, os preços refletem completamente todas as informações disponíveis, é considerado "eficiente". Neste paradigma, Fama também apresenta um modelo no qual as mudanças sucessivas de preço são independentes e distribuídas de forma idêntica, chamado de modelo de Random Walk. Ambos os conceitos apoiam a ideia de que, uma vez que todas as informações disponíveis já estão refletidas no mercado, a previsão de preço de qualquer ativo é inútil.

2.2 Quebra-Cabeça de Meese e Rogoff

Meese e Rogoff (1983a) compararam três modelos teóricos: o modelo monetário de preço flexível, o modelo monetário de preço rígido e o modelo de Hooper-Morton; para os preços do dólar americano em relação ao iene, à libra e ao marco alemão; e descobriram que nenhum dos modelos alcança um erro quadrático médio (RMSE) menor do que o modelo de Random Walk em um horizonte de um a doze meses. Em seu trabalho subsequente, Meese e Rogoff (1983b) afirmam que o fraco desempenho desses modelos teóricos não é devido a inconsistência nos parâmetros, porque mesmo ao aplicar técnicas autorregressivas, esses modelos têm desempenho inferior ao modelo de Random Walk. Eles concluem que, mesmo com técnicas mais eficientes, há uma baixa chance de que esse resultado mude.

2.3 Imprevisibilidade Fundamental

Engel et al. (2007) afirmam que os resultados de Meese e Rogoff não implicam que esses modelos tenham sido refutados pelos dados. Na verdade, eles encontraram evidências de que os modelos teóricos ajudam a prever mudanças quando se utilizam modelos em painel. Seus resultados mostraram que o modelo de painel geralmente produz uma previsão melhor do que o modelo de Random Walk em horizontes longos. Essa divergência de resultados entre os trabalhos de Meese e Rogoff e Engel, é abordada na revisão da literatura de Rossi (2013). Em

sua revisão sistemática dos estudos mais relevantes sobre a previsão de taxa de câmbio, Rossi explica que a divergência de resultados presentes na literatura, dependente do horizonte e da moeda estudados. Alguns modelos apresentam menores erros em horizontes curtos, enquanto outros em horizontes longos, porém, nenhum deles é capaz de superar o modelo de random walk em todos os períodos.

Apesar desses modelos apresentarem baixa capacidade de previsão, Rossi (2005) afirma que os modelos econométricos fornecem uma melhor descrição dos dados do que o modelo de Random Walk. De maneira semelhante, Moosa e Burns (2013) explicam que estes modelos não conseguem superar o modelo de random walk em termos de RMSE como métrica de previsão, porém são melhores ao apontar a direção e a magnitude do movimento de preços. Bacchetta et al. (2009) propõem que o desempenho inferior dos modelos econômicos se deve à relação instável entre a taxa de câmbio e os fundamentos. Essa relação aumenta os erros de estimação dos parâmetros, implicando que o comportamento passado é improvável de ser útil para prever o futuro. Além disso, eles descobriram que, ao adicionar parâmetros variáveis no tempo, o efeito líquido no desempenho da previsão fora da amostra é inexistente.

2.4 Machine Learning

Mesmo como um histórico de resultados negativos, o desenvolvimento de tecnologias recentes permitiu que pesquisadores explorassem metodologias alternativas para a previsão da taxa de câmbio. A tecnologia mais promissora para esse fim é o Machine Learning (ML), que, diferente da programação tradicional, onde se faz necessário o desenvolvimento de algoritmos a serem executado por um computador, permite que este aprenda dinamicamente como realizar tarefas e identificar padrões Russel e Norig (2021). Dentro do escopo de machine learning, existem metodologias robustas desenvolvidas para resolver problemas complexos chamadas de Deep Learning. Russel e Norig (2021) definem Deep Learning como técnicas de aprendizado nas quais hipóteses são traduzidas em circuitos complexos fortemente conectados. Modelos que utilizam deep learning são capazes de treinar longos caminhos computacionais com interações abstratas e são mais adequados para modelos de previsão de séries temporais.

Dentro do escopo do Deep Learning, também encontramos limitações quando utilizamos redes muito longas e complexas. Isto foi confirmado no estudo feito por Bengio et al. (1994) onde verificou-se que, ao treinar uma rede neural recorrente, à medida que o alcance temporal aumenta, o gradient descent se torna cada vez mais ineficiente, o modelo é incapaz de reter a informação fornecida ao longo de toda a cadeia aumentando seu erro consideravelmente. Como solução deste problema de retroalimentação do erro presente nos primeiros modelos de Deep Learning, Hochreiter e Schmidhuber (1997) propuseram uma nova arquitetura de rede recorrente com algoritmo de aprendizado baseado em gradiente chamado Long Short-Term Memory (LSTM). Este novo algoritmo, tornou mais viável o uso de Deep Learning para treinamento de grande volume de dados, permitindo que pesquisadores o utilizem como uma alternativa aos modelos econômicos tradicionais para prever o movimento cambial, e serve como base para outros modelos de deep learning. Neste estudo, busca-se superar o puzzle de Meese e Rogoff utilizando diferentes tipos de técnicas de machine learning, como o LSTM, e outros tipos de modelos influenciados por essa metodologia.

3 METODOLOGIA

Nesta seção, apresenta-se as amostras e métodos utilizados para a criação dos modelos estudados.

3.1 Amostras

Em busca de testar a eficácia do modelo em diferentes mercados, se utilizarão 5 moedas para comparação contra o dólar americano: o real brasileiro, o euro, o yuan chinês, a libra esterlina e o iene japonês. Para cada uma das moedas estudadas, foram extraídos dados diários entre janeiro de 2010 a dezembro de 2023 utilizando a plataforma Bloomberg.

Tabela 1 - Análise Descritiva

Variables	Samples	Mean	Std. deviation	Min	Max	Kurtosis	Skewness	Variance
<i>USDBRL</i>	3,260	3.570	1.278	1.539	5.887	-0.413	0.495	0.022
<i>USDCNH</i>	3,482	6.579	0.325	6.020	7.365	-0.856	0.364	0.105
<i>USDEUR</i>	3,392	0.849	0.076	0.674	1.042	-0.828	-0.311	0.006
<i>USDGBP</i>	3,393	0.718	0.077	0.583	0.936	-1.254	-0.033	0.006
<i>USDJPY</i>	3,393	108.683	17.078	75.820	151.720	0.110	0.103	291.670

Fonte: Elaboração Própria

3.2 Modelos de Previsão baseados em Machine Learning

O aprendizado de máquinas pode ser compreendido como “o descobrimento de um grupo de valores para os pesos de todas as camadas de uma rede neural, de forma que esta rede consiga mapear corretamente os inputs para os alvos desejados” (Chollet, 2021). A Tabela 2 mostra os parâmetros (pesos) para as cinco técnicas de machine learning aplicadas nos desenvolvimentos dos modelos de previsão cambial.

Tabela 2 - Parâmetros por Modelo

		USDBRL	USDCNH	USDEUR	USDGBP	USDJPY
LSTM	<i>Layers</i>	4	4	4	4	4
	<i>Learning Rate</i>	0.01	0.001	0.000001	0.000001	0.001
	<i>Epochs</i>	2000	2000	2000	2000	2000
GAN	<i>Layers</i>	128	128	128	128	128
	<i>Learning Rate</i>	0.01	0.01	0.01	0.01	0.01
	<i>Epochs</i>	200	200	200	200	200
N-BEATS	<i>Layers</i>	4	4	4	4	4
	<i>Learning Rate</i>	0.001	0.001	0.001	0.001	0.001
	<i>Epochs</i>	100	100	100	100	100
N-HITS	<i>Layers</i>	2	2	2	2	2
	<i>Learning Rate</i>	0.001	0.001	0.001	0.001	0.001
	<i>Epochs</i>	200	200	200	200	200

Fonte: Elaboração Própria

3.3 Long Short-Term Memory (LSTM)

A primeira técnica é baseada na abordagem de LSTM. Esse tipo de modelo pode cobrir defasagens temporais ao longo de mais de 1000 passos discretos de tempo, ao impor um erro constante dentro de unidades especiais (Hochreiter & Schmidhuber, 1997), o que resolve os problemas clássicos de desaparecimento em modelos de machine learning anteriores e permite seu uso para previsão de dados de séries temporais.

3.4 Generative Adversarial Network

A segunda técnica é a Generative Adversarial Network (GAN). Esse modelo funciona com base em um jogo entre duas redes: o gerador e o discriminador (Goodfellow et al., 2020). O primeiro tenta replicar as amostras de treinamento a partir de um ruído aleatório, enquanto o último compara as duas para encontrar qual é a amostra de treinamento e qual é a falsa criada pela rede geradora. Em algum momento, o gerador cria cópias tão perfeitas que o discriminador não consegue distingui-las.

3.5 Neural Basis Expansion Analysis for Interpretable Time Series Forecasting

A terceira técnica é chamada de Neural Basis Expansion Analysis for Interpretable Time Series Forecasting (N-BEATS). É uma arquitetura inovadora criada para resolver o problema de previsão de pontos de séries temporais univariadas, com base em links residuais anteriores e posteriores em camadas profundas totalmente conectadas (Oreshkin et al., 2020).

3.6 Neural Hierarchical Interpolation for Time Series Forecasting (N-HiTs)

A quarta técnica é chamada de Neural Hierarchical Interpolation for Time Series Forecasting (N-HiTs). É um novo algoritmo de previsão neural composto pela combinação de amostragem de entrada de múltiplas taxas e interpolação hierárquica (Challu et al., 2022). Este modelo aumenta o desempenho da previsão da taxa de câmbio em longos horizontes.

3.7 Prophet

A quinta técnica é chamada de Prophet, desenvolvido pelo Facebook formado por dois componentes: o primeiro contém um modelo de regressão modular que permite ao usuário selecionar os componentes relevantes para o processo de previsão, o segundo é um sistema para medir e avaliar a precisão do modelo, e criar alertas caso haja necessidade de ajustes manuais (Taylor e Letham, 2018).

4 ANÁLISE DOS RESULTADOS

Tabela 3 - RMSE por Modelo

Models	USDBRL	USDCNH	USDEUR	USDGBP	USDJPY
<i>ARIMA</i>	0.269	0.353	0.034	0.020	7.673
<i>GAN</i>	3.397	6.249	0.417	0.359	109.810
<i>LSTM</i>	0.050	0.020	0.035	0.052	0.770
<i>N-BEATS</i>	0.084	0.111	0.038	0.016	1.088
<i>N-HiTs</i>	0.535	0.301	0.112	0.047	19.007
<i>Prophet</i>	0.363	0.640	0.034	0.061	13.015
<i>Random Walk</i>	0.016	0.011	0.005	0.003	0.530

Fonte: Elaboração Própria

A Tabela 3 mostra que nenhum dos modelos de machine learning alcançou um RMSE menor do que o modelo de Random Walk, no entanto, os modelos LSTM e N-BEATS exibiram valores de RMSE próximos ao desejado, e acima da performance do modelo ARIMA. Isto sugere que, estes modelos são os mais eficientes para a previsão de dados de câmbio, além disso, avanços em ambas técnicas ou ajustes minuciosos dos parâmetros podem resultar em uma melhor performance em estudos futuros.

Por outro lado, o modelo GAN apresentou a pior performance, indicando que esta técnica não é adequada para este tipo de tarefa e, a menos que haja novos avanços, ela deve ser abandonada como ferramenta de predição cambial. Os modelos N-HiTs e Prophet apresentaram performances similares, porém abaixo do esperado, apesar possuírem um potencial em estudos futuros, os resultados sugerem que existem modelos melhores a serem explorados.

Ainda que parte dessas descobertas sejam promissoras, o puzzle de Meese e Rogoff se sustenta, reafirmando a hipótese de que todas as informações disponíveis estão totalmente refletidas nos preços dos ativos, e, portanto, é impossível prever seus valores em uma economia eficiente.

5 CONCLUSÃO

O puzzle de Meese e Rogoff é o principal desafio na literatura sobre previsão da taxa de câmbio, este estudo introduz 5 técnicas de predição de dados de machine learning com o objetivo de superar esse desafio. Apesar de nenhum dos modelos apresentados produzirem uma performance melhor que o modelo de Random Walk, os resultados sugerem que os modelos LSTM e N-BEATS representam alternativas viáveis para este fim. A performance de ambos os modelos, superou o modelo ARIMA para a maior parte das moedas, e seus valores se aproximaram da performance do modelo de Random Walk, indicando que dos modelos estudados, os dois modelos são os mais adequados para a previsão do câmbio. Em estudos futuros, pode-se aprofundar o uso desses dois modelos alterando seus parâmetros e quantidade de observações analisadas para a obtenção de uma melhor performance. Além disso, este tipo de tecnologia tem crescido exponencialmente, em um futuro breve essas técnicas podem ser aprimoradas permitindo resultados mais satisfatórios.

REFERÊNCIAS

- BACCHETTA, P.; VAN WINCOOP, E.; BEUTLER, T. Can Parameter Instability Explain the Meese-Rogoff Puzzle? **NBER International Seminar on Macroeconomics**, v. 6, n. 1, p. 125–173, 2009. Disponível em: <https://www.jstor.org/stable/10.1086/648702>.
- BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning Long-Term Dependencies with Gradient Descent is Difficult. **IEEE Transactions On Neural Networks**, v. 5, n. 2, p. 157–166, 1994. Disponível em: <https://doi.org/10.1109/72.279181>. Acesso em: 18 maio 2024.
- CHALLU, C.; OLIVARES, K. G.; ORESHKIN, B. N.; RAMIREZ, F. G.; CANSECO, M. M.; DUBRAWSKI, A. Nhits: Neural hierarchical interpolation for time series forecasting. Disponível em: **PROCEEDINGS of the AAI Conference on Artificial Intelligence**, v. 37, n. 6, p. 6989-6997, jun. 2023.
- CHOLLET, Francois. **Deep learning with Python**. 2. ed. Simon and Schuster, 2021.
- ENGEL, C. et al. Exchange Rate Models Are Not as Bad as You Think. **National Bureau of Economic Research**, v. 22, p. 381-441, 443-473, 2007.
- FAMA, E. F. Efficient capital markets. **Journal of finance**, v. 25, n. 2, p. 383–417, 1970.
- GERS, F. A.; SCHMIDHUBER, J.; CUMMINS, F. Learning to Forget: Continual Prediction with LSTM. **Neural Computation**, v. 12, n. 10, p. 2451–2471, 2000. Disponível em: <https://doi.org/10.1162/089976600300015015>.
- GOODFELLOW, I. et al. Generative Adversarial Networks. **Communications of the ACM**, v. 63, n. 11, 2020. Disponível em: <https://doi.org/10.1145/3422622>.
- HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997. Disponível em: <https://doi.org/10.1162/neco.1997.9.8.1735>.
- MESEE, R. A.; ROGOFF, K. Empirical Exchange Rate Models Of The Seventies: Do They Fit Out Of Sample? **Journal of International Economics**, v. 14, p. 3–24, 1983.
- MEESE, R. A.; ROGOFF, K. The Out-Of-Sample Failure Of Empirical Exchange Rate Models: Sample Error Or Misspecification? **International Finance Discussion Papers**, n. 42, 1983.
- MOOSA, I.; BURNS, K. A reappraisal of the Meese–Rogoff puzzle. **Applied Economics**, v. 46, n. 1, p. 30–40, 2013. Disponível em: <https://doi.org/10.1080/00036846.2013.829202>.
- ORESHKIN, B. N.; CARPOV, D.; CHAPADOS, N.; BENGIO, Y. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. **arXiv preprint arXiv:1905.10437**, 2019.
- ROSSI, B. Testing Long-Horizon Predictive Ability With High Persistence, And The Meese–Rogoff Puzzle. **International Economic Review**, v. 46, n. 1, p. 61–92, 2005.
- ROSSI, B. **Exchange Rate Predictability**. Barcelona GSE Working Paper Series, n. 690, 2013.
- RUSSEL, S.; NORVIG, P. Artificial Intelligence: A Modern Approach. 4. ed. **Pearson Education**, 2021.
- TAYLOR, Sean J.; LETHAM, Benjamin. Forecasting at scale. **The American Statistician**, v. 72, n. 1, p. 37-45, 2018.