INSPER: INSTITUTO DE ENSINO E PESQUISA
PROGRAMA DE MESTRADO PROFISSIONAL EM ECONOMIA

BRUNO PAULO FERREIRA

PREVISÃO DA INFLAÇÃO NO BRASIL UTILIZANDO ABORDAGEM DESAGREGADA E APRENDIZADO DE MÁQUINA

SÃO PAULO - SP

2023

BRUNO PAULO FERREIRA

PREVISÃO DE INFLAÇÃO NO BRASIL UTILIZANDO A ABORDAGEM DESAGREGADA E APRENDIZADO DE MÁQUINA

Dissertação apresentada ao Programa de Mestrado Profissional em Economia do Insper – Instituto de Ensino e Pesquisa como requisito parcial para a obtenção do título de Mestre em Economia.

Área de concentração: Macroeconomia, Machine Leaning.

Orientador: Prof. Dr. Miguel Bandeira

SÃO PAULO - SP

2023

Paulo Ferreira, Bruno

Previsão da Inflação no Brasil usando Abordagem Desagregada e Aprendizado de Máquina. /Bruno Paulo Ferreira. – São Paulo, 2023.

Dissertação (Mestrado Profissional em Economia) - Insper – Instituto de Ensino e Pesquisa como requisito parcial para a obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. Miguel Bandeira

1. Macroeconomia. 2. Machine Learning. 3. Econometria. 4. Inflação. 5. Previsão Inflação.

BRUNO PAULO FERREIRA

PREVISÃO DE INFLAÇÃO NO BRASIL UTILIZANDO ABORDAGEM DESAGREGADA E APRENDIZADO DE MÁQUINA

Dissertação apresentada ao Programa de Mestrado Profissional em Economia do Insper – Instituto de Ensino e Pesquisa como requisito parcial para a obtenção do título de Mestre em Economia.

Área de concentração: Macroeconomia, Inflação.

Orientador: Prof. Dr. Miguel Bandeira

Data de Aprovação: ____/____/____

_____

Prof. Dr. Diogo Abry Guillen

Banco Central do Brasil

_____

Prof. Dr. Gustavo Silva Araujo

FGV-EPGE

SÃO PAULO - SP

2023

Resumo

O presente trabalho aborda a utilização de técnicas de aprendizado de máquina para prever a inflação agregada a partir de dados desagregados de preços de diferentes categorias. São utilizados dados de séries temporais de preços de bens e serviços, bem como dados macroeconômicos relevantes, para construir o modelo de previsão. O modelo é construído usando uma variedade de algoritmos de aprendizado de máquina, incluindo modelos de regularização, árvores de decisão e máquinas de vetor de suporte (SVM). Os resultados indicam que o modelo proposto apresenta desempenho superior aos modelos de previsão adotados como benchmark neste estudo. Além disso, o estudo mostra que o desempenho do modelo é aprimorado quando os dados desagregados são incluídos no modelo.

Palavras-chave: 1. Macroeconomia. 2. Machine Learning. 3. Econometria. 4. Inflação. 5. Previsão Inflação.

Abstract

This study addresses the use of machine learning techniques to predict aggregate inflation from disaggregated data of prices from different categories. Time series data of prices of goods and services, as well as relevant macroeconomic data, are used to build the prediction model. The model is constructed using a variety of machine learning algorithms, including regularization models, decision trees, and support vector machines (SVM). The results indicate that the proposed model performs better than the prediction models adopted as benchmarks in this study. Furthermore, the study shows that the model's performance is improved when disaggregated data is included in the model.

Keyword: 1. Macroeconomics. 2. Machine Learning. 3. Econometrics 4. Inflation 5. Inflation Forecast

SUMÁRIO

1. INTRODUÇÃO 7
1.1 REVISÃO DA LITERATURA 8
1.2 ORGANIZAÇÃO DESTE ESTUDO 10
1.3 CONTRIBUIÇÃO PARA A LITERATURA 10
2. METODOLOGIA 11
2.1 ESTACIONARIEDADE 11
2.2 VALIDAÇÃO CRUZADA 11
2.3 CRITÉRIO DE AVALIAÇÃO 13
3. BASE DE DADOS 11
3.1 DESAGREGAÇÃO 11
3.2 VARIÁVEIS EXPLICATIVAS 11
4. MODELOS 17
5. RESULTADOS 24
6. CONCLUSÕES 34
REFERÊNCIAS 35

1. INTRODUÇÃO

Em junho de 1999, o Brasil implementou o Sistema de Metas para Inflação como um meio de combater a inflação, que historicamente tem sido persistentemente alta no país. O objetivo principal do sistema era estabilizar a economia brasileira e fornecer uma base para a tomada de decisões econômicas, incluindo as políticas monetária, fiscal e cambial. A política monetária define a taxa de juros básica da economia, que é capaz de influenciar o desemprego, a atividade econômica e a inflação. O objetivo principal do sistema de metas é garantir que a inflação permaneça dentro da meta definida pelo CMN (Conselho Monetário Nacional), que também estabelece as bandas superior e inferior da meta. O índice de preços utilizado para definir a meta é o IPCA (Índice de Preços ao Consumidor Amplo), medido pelo IBGE (Instituto Brasileiro de Geografia e Estatística), e é considerado o principal índice de preços do país. Portanto, previsões precisas de inflação são importantes para a economia como um todo e são cruciais para que o Banco Central tome decisões assertivas sobre a política monetária.

Prever a inflação é essencial para diversas decisões econômicas, tais como investimentos, fixação de preços, decisões das empresas, negociação de salários e outros contratos, bem como para a condução de políticas monetárias e fiscais. Diante disso, economistas e econometristas buscam constantemente ferramentas cada vez mais precisas e eficientes para fornecer previsões da inflação. Com a grande quantidade de informações disponíveis atualmente, novas abordagens estatísticas e econométricas têm surgido para lidar com essa questão antiga. Além de considerar defasagens na própria inflação e várias variáveis macroeconômicas como preditores, é interessante observar as desagregações do índice de preços (mais especificamente, suas variações), uma vez que essa abordagem pode levar a um desempenho preditivo melhor para a inflação agregada. Portanto, previsões confiáveis de inflação são cruciais para a economia como um todo e são fundamentais para a tomada de decisões informadas sobre a política monetária.

Este estudo tem como objetivo comparar o desempenho da previsão de inflação ao agregar previsões desagregadas com abordagens mais convencionais presentes na literatura, tais como previsões baseadas em pesquisa e previsões feitas diretamente a partir do agregado. Considerando múltiplas desagregações, realizamos exercícios de previsão utilizando tanto abordagens de previsão desagregada quanto de previsão agregada. Ao levar em conta o aumento da granularidade informacional que as desagregações proporcionam, elas podem fornecer uma potencial vantagem no desempenho da previsão de inflação. Além disso, utilizamos técnicas de

séries temporais tradicionais e de aprendizado de máquina linear e não-linear para lidar com um grande número de preditores, as quais são amplamente utilizadas na pesquisa econômica e têm se mostrado ferramentas valiosas.

A utilização de séries temporais desagregadas possibilita a análise da dinâmica de preços tanto agregados quanto desagregados, bem como a investigação de diferentes abordagens de previsão. O estudo de Bermingham and D'Agostino (2014) ressalta a importância de considerar a heterogeneidade dos dados ao utilizar séries temporais desagregadas, o que pode aprimorar a capacidade dos modelos econométricos em capturar a dinâmica de preços.

1.1 REVISÃO DA LITERATURA

Apesar da importância da previsão de inflação em diversos aspectos, não há unanimidade na literatura sobre qual é a melhor abordagem econométrica para realizar previsões precisas. Isso se deve, em grande parte, às propriedades das séries temporais da inflação, que são reconhecidas por apresentar características não-estacionárias e não-lineares.

A análise empírica realizada por Stock e Watson (1999) comparou diferentes modelos de previsão de inflação, incluindo modelos univariados, modelos que utilizam a curva de Phillips e modelos que usam um conjunto de variáveis macroeconômicas como inputs. Os resultados indicaram que os modelos que usam variáveis macroeconômicas como inputs apresentam um desempenho superior aos outros modelos na previsão da inflação. Atkeson e Ohanian (2001) também destacaram que as curvas de Phillips não são confiáveis para prever a inflação devido às relações entre o desemprego e a inflação mudarem ao longo do tempo. Eles argumentaram que modelos que incorporam expectativas e consideram choques estruturais são mais adequados para prever a inflação. Andrew, Bekaert & Wei (2007) compararam o desempenho de três abordagens para prever a inflação: (i) modelos baseados em variáveis macroeconômicas, (ii) preços de ativos financeiros e (iii) pesquisas de expectativas de inflação. Os autores concluíram que as pesquisas de expectativas de inflação superam os outros dois métodos em termos de precisão na previsão de inflação no curto prazo. No entanto, no longo prazo, os modelos baseados em variáveis macroeconômicas tendem a ter um melhor desempenho.

Inoue e Kilian (2008) investigaram a eficácia do método de bootstrap aggregating (bagging) na previsão da inflação nos Estados Unidos, comparando-o com outros modelos de previsão, como ARIMA, redes neurais e modelos de fatores. Os resultados sugerem que o método de bagging pode melhorar significativamente a precisão das previsões em relação a outros modelos, especialmente em horizontes de previsão mais longos. Por outro lado, Monteforte e Moretti (2013) utilizaram uma abordagem de agrupamento (clustering) para agrupar os ativos financeiros em diferentes categorias e, em seguida, usaram essas categorias como variáveis explicativas em um modelo de regressão para prever a inflação. O modelo proposto foi comparado com modelos tradicionais de previsão de inflação e os resultados mostram que o modelo de agrupamento supera os modelos tradicionais em termos de precisão das previsões.

Medeiros, Garcia e Vasconcelos (2017) abordam a previsão da inflação brasileira usando modelos de alta dimensionalidade. Eles comparam a precisão das previsões de inflação usando modelos de fatores dinâmicos, modelos Bayesianos de redes neurais e modelos de árvores de regressão. Os resultados indicam que os modelos de fatores dinâmicos e Bayesianos de redes neurais superam as previsões dos modelos de referência, especialmente em horizontes mais longos. Os autores também discutem a relevância de modelos de alta dimensionalidade para a previsão da inflação, em particular em países emergentes como o Brasil, onde a complexidade e a volatilidade da economia tornam a previsão da inflação uma tarefa desafiadora.

Theoharidis (2021) avançou na área de Deep Learning e propõe o uso de redes neurais convolucionais e recorrentes com memória de longo prazo (LSTM) para prever a inflação em países dos Estados Unidos. Além disso, o autor também utiliza variational autoencoders (VAE) para reduzir a dimensionalidade dos dados e melhorar a precisão das previsões. Os resultados mostram que as redes neurais convolucionais e LSTM superam os modelos tradicionais de previsão de inflação, e que a combinação de VAE e LSTM é capaz de melhorar ainda mais a precisão das previsões.

Florido (2021) apresenta o uso de modelos de aprendizado de máquina para prever a inflação no Brasil, tanto em termos agregados quanto desagregados por categorias de produtos. Os resultados indicam que os modelos de aprendizado de máquina superam os modelos tradicionais na previsão da inflação, especialmente em horizontes de curto e médio prazo. Boaretto e Medeiros (2022) seguem uma linha de pesquisa semelhante, abordando a previsão da inflação agregada com dados desagregados e técnicas de aprendizado de máquina. Eles

comparam modelos baseados em dados agregados com modelos que usam dados desagregados de componentes da inflação e examinam a eficácia de diferentes algoritmos de aprendizado de máquina na previsão da inflação. Os resultados indicam que o uso de dados desagregados e técnicas de aprendizado de máquina pode melhorar significativamente a precisão da previsão da inflação agregada.

1.2 ORGANIZAÇÃO DESTE ESTUDO

Este artigo tem outras cinco seções além da introdução. A seção 2 explica a metodologia utilizada e apresenta três fatores que ajudam a definir questões fundamentais: i) estacionariedade; ii) validação cruzada; iii) critério de avaliação. O primeiro item é importante porque várias variáveis financeiras usadas neste estudo são não estacionárias, e alguns modelos usados neste estudo exigem que as variáveis sejam estacionárias. A validação cruzada é usada para evitar o possível overfitting gerado pelos modelos de aprendizado de máquina. Essa técnica envolve a divisão dos dados em blocos de treinamento e teste para determinar qual modelo produz o menor erro de previsão. O critério de avaliação é usado para classificar o melhor modelo. Existem várias opções usadas na literatura, e escolhemos uma métrica amplamente utilizada: MAE (Erro Absoluto Médio).

Na seção 3, são utilizadas diversas variáveis financeiras e macroeconômicas como explicativas para otimizar os modelos. Na seção 4, o modelo ARIMA é apresentado como referência do estudo, juntamente com vários outros modelos categorizados como Machine Learning, como modelos de regularização, árvores de decisão e máquinas de vetor de suporte. Por fim, na seção 5, são apresentados os resultados dos melhores modelos para cada desagregação e o resultado consolidado.

1.3 CONTRIBUIÇÃO PARA A LITERATURA

Este artigo é pioneiro na aplicação do conceito de blocos precedentes móveis como método de validação cruzada nos modelos de previsão de inflação no Brasil. A validação cruzada com blocos precedentes móveis demonstrou ter ótimos resultados em Cerqueira et al. 2019 e é uma técnica que se destaca por reduzir o tempo de otimização de cada modelo (Cerqueira et al. 2021).

Em segundo lugar, este artigo propôs diversas combinações de modelos de aprendizado de máquina com o objetivo de descobrir se a maior contribuição vem da desagregação ou da utilização dos modelos. Desta forma, esses resultados oferecem uma valiosa contribuição para a macroeconomia aplicada, principalmente na era do Big Data e especialmente no que diz respeito à previsão de inflação no Brasil.

2. METODOLOGIA

2.1 ESTACIONARIEDADE

Para garantir a estacionaridade de cada variável explicativa, foi realizada a transformação de séries temporais seguindo o procedimento adotado por McCracken e Ng (2016). Esse procedimento propõe seis tipos de transformações: (1) sem tranformação ; (2) ΔX_t ; (3) $\Delta^2 X_t$; (4) $\ln(X_t)$; (5) $\Delta \ln(X_t)$; and (6) $\Delta^2 \ln(X_t)$.

Após a realização das transformações nas séries temporais, foi utilizado o teste KPSS (Kwiatkowski-Phillips-Schmidt-Shin) para verificar a estacionaridade de cada série, adotando um nível de significância de 5%.

2.2 VALIDAÇÃO CRUZADA

Validação cruzada é uma técnica crucial na avaliação da capacidade de generalização de um modelo, evitando tanto o underfitting quanto o overfitting. Essa técnica consiste na repetição do experimento com diversas subamostras da amostra original, permitindo que o modelo seja testado em diferentes conjuntos de dados e sua capacidade de generalização seja avaliada de forma mais precisa. Assim, a validação cruzada é uma ferramenta fundamental para garantir a confiabilidade e robustez dos modelos de aprendizado de máquina.

2.2.1 VALIDAÇÃO CRUZADA RANDÔMICA

A validação cruzada randômica é uma técnica de validação cruzada amplamente utilizada em diferentes áreas de pesquisa. Ela permite avaliar o desempenho do modelo de forma mais robusta, garantindo que a avaliação não seja influenciada pela divisão arbitrária dos dados em conjuntos de treinamento e teste.

No entanto, a validação cruzada randômica pode ser problemática em séries temporais, já que os dados geralmente apresentam uma estrutura temporal e não são independentes e identicamente distribuídos (i.i.d.). Para contornar esse problema, uma abordagem comumente utilizada é a técnica de blocos precedentes móveis, que consiste em dividir os dados em blocos de tamanho fixo e usar cada bloco como um conjunto de teste, com os blocos anteriores usados como conjuntos de treinamento.

Essa abordagem leva em conta a estrutura temporal dos dados, garantindo que o modelo seja avaliado em conjuntos de teste que sejam compatíveis com a série temporal. Além disso, a técnica de blocos precedentes móveis pode ser computacionalmente mais eficiente do que a validação cruzada randômica, pois requer menos divisões aleatórias dos dados.

2.2.2 BLOCOS PRECEDENTES MÓVEIS (PREQUENTIAL SLIDING BLOCKS)

Prequential sliding blocks, também conhecidos como blocos precedentes móveis, é uma técnica de validação cruzada desenvolvida especialmente para séries temporais. A ideia é dividir a série temporal em blocos de tamanho fixo e usar cada bloco como conjunto de teste, enquanto os blocos anteriores são usados como conjunto de treinamento.

Essa técnica é vantajosa porque leva em consideração a estrutura autoregressiva da série temporal, evitando o problema de quebrar essa estrutura, como pode ocorrer na validação cruzada randômica. Além disso, a técnica é capaz de detectar mudanças nas propriedades da série temporal ao longo do tempo, o que a torna mais adequada para séries com não-estacionariedades e outras complexidades temporais.

Para implementar essa técnica, é necessário escolher um tamanho de bloco apropriado para a série temporal em questão e manter fixo o tamanho do conjunto de treinamento. Em seguida, desliza-se o bloco pelos dados para criar subconjuntos de treinamento/teste. É importante ressaltar que a técnica pode ser computacionalmente intensiva, principalmente se o tamanho do bloco for grande.

Então, em geral, a técnica de blocos precedentes móveis tem obtido ótimo desempenho em estudos de séries temporais e pode ser uma escolha apropriada para avaliar modelos que lidam com esses dados. Cerqueira et al. (2019) obteve resultados positivos ao testar essa técnica em séries temporais, e posteriormente, Cerqueira et al. (2021) destacou o tempo de processamento como um quesito em que essa técnica se mostrou eficiente.

2.3 CRITÉRIO DE AVALIAÇÃO

O critério de avaliação é fundamental para verificar a qualidade de um modelo de machine learning. Ele permite comparar as previsões feitas pelo modelo com os valores reais e determinar o quão bem o modelo está se ajustando aos dados. Existem diversas medidas de avaliação disponíveis, que variam de acordo com o tipo de problema e objetivo do modelo. Algumas das medidas mais comuns são: (i) Erro médio absoluto (MAE), que representa a média das diferenças absolutas entre as previsões e os valores reais; (ii) Erro quadrático médio (MSE), que representa a média das diferenças quadráticas entre as previsões e os valores reais; e (iii) Raiz do erro quadrático médio (RMSE), que é a raiz quadrada do MSE.

Para verificar a precisão do modelo, é importante testá-lo com dados reais. Para evitar o overfitting, é comum separar a amostra em dados de treinamento e de teste. Quando os erros de treinamento são muito maiores do que os de teste, o modelo pode estar sofrendo de underfitting. Por outro lado, quando os erros de treinamento são bem menores do que os de teste, o modelo pode estar sofrendo de overfitting.

Para escolher o melhor modelo de previsão, é comum utilizar uma medida de avaliação como critério. No seu caso, escolhemos o MAE (Mean Absolute Error), com base no estudo de Willmott & Matsuura (2005). Este artigo discute a escolha de uma medida de avaliação adequada para avaliar o desempenho de modelos de previsão climática. Os autores comparam a utilização do Erro Médio Absoluto (MAE) e do Erro Quadrático Médio (RMSE) como medidas de avaliação. Eles argumentam que o MAE é uma medida mais intuitiva e fácil de interpretar do que o RMSE, uma vez que leva em conta a magnitude do erro, independentemente da sua direção. Além disso, eles mostram que o MAE é mais robusto em relação a outliers do que o RMSE. Os autores concluem que o MAE é uma medida mais apropriada para avaliar modelos de previsão climática em comparação com o RMSE.

3. BASE DE DADOS

Este estudo tem como objetivo analisar e comparar diferentes modelos de previsão para o IPCA, que é o índice de preços ao consumidor amplo medido pelo Instituto Brasileiro de Geografia e Estatística (IBGE). O IPCA é utilizado para calcular a inflação oficial no Brasil e, portanto, é um indicador importante para a política monetária do país. O estudo utiliza horizontes de previsão que variam de 1 mês até 18 meses, e a amostra tem início em agosto de

2006 e término em agosto de 2022, totalizando 193 observações. Ao comparar diferentes modelos, o estudo visa identificar o modelo que apresenta a melhor performance em termos de precisão das previsões.

3.1 IPCA (DESAGREGAÇÃO)

O Instituto Brasileiro de Geografia e Estatística (IBGE) divide o IPCA em diversos grupos e subgrupos. O objetivo deste trabalho é comparar diversos modelos para previsão do IPCA fora da amostra. Desta forma, iremos propor algumas agregações [1]: i) agregação 1 (Agg 1): número cheio (headline); ii) agregação 2 (Agg 2): administrados + livres; iii) agregação 3 (Agg 3): administrados + alimentação no domicílio + serviços + industriais; iv) agregação 4 (Agg 4): administrados + tradables + nontradables.

Tabela 1

	Agg1	Agg2	Agg3	Agg4
IPCA	1	-	-	-
Administrados	-	1	1	1
Livres (por atividade econômica)	-	1	-	-
Alimentação no Domicílio	-	-	1	-
Serviços	-	-	1	-
Industriais	-	-	1	-
Livres (por tipo de produto)		-	-	-
Tradables	-	-	-	1
Non-Tradables	-	-	-	1

Fonte: Florido (2021)

[1] Os pesos de cada categoria foram baseados na Pesquisa de Orçamentos Familiares (POF) realizada em 2018.

3.2 VARIÁVEIS EXPLICATIVAS

Neste estudo, foram utilizadas 67 variáveis financeiras e macroeconômicas listadas na tabela abaixo, com o objetivo de prever o IPCA fora da amostra. Para evitar o problema de overfitting, foram utilizados 3 lags das variáveis. Além disso, foram adicionadas 11 dummies sazonais para lidar com o forte componente sazonal presente na inflação brasileira.

Tabela 2

Série	Fonte
Índice Nacional de Preços ao Consumidor (INPC)	BCB
Índide Geral de Preços - Mercado (IGP-M) - FGV	BCB
Índide Geral de Preços - Disponibilidade Interna (IGP-DI) - FGV	BCB
IPC-BR	BCB
Índice Nacional de Custo de Construção (INCC)	BCB
Índice de Preços ao Consumidor - São Paulo (IPC-Fipe)	BCB
Índice Geral de Preços - 10 (IGP-10)	BCB
IPCA-15	BCB
Primário (Ajustado Sazonalmente c/ X13)	BCB
Taxa de Câmbio	BCB
Focus - Inflação 12m Suavizada	BCB
Meios de pagamento amplos - M1 - Saldo em final de período	BCB
Meios de pagamento amplos - M2 - Saldo em final de período	BCB
Meios de pagamento amplos - M3 - Saldo em final de período	BCB
Meios de pagamento amplos - M4 - Saldo em final de período	BCB
Taxa de Juros de Longo Prazo (TJLP)	BCB
Taxa Selic	BCB
PIB Mensal - valores correntes	BCB
Dívida Líquida do Setor Público - Total - Setor Público Consolidado	BCB

Resultado Primário - Total - Setor Público Consolidado	BCB
Vendas de Veículos pelas concessionáris - Automóveis - Fenabrave	BCB
Balança Comercial - Balanço de Pagamentos BPM6 - Mensal - Saldo	BCB
IBC-Br (com ajuste sazonal)	BCB
Indicadores de Produção Industrial - Geral - Dados Dessazonalizados	BCB
Tráfego de Veículos Pesados - ABCR - Daddos Dessazonalizados	BCB
Consumo de Energia Elétrica - Comercial	BCB
Utilização da Capacidade Instalada - Indústria de Transformação (CNI) - SA	BCB
Utilização da Capacidade Instalada - Indústria de Transformação (FGV) - SA	BCB
Estoque de Empregos Formais - Total - CAGED	BCB
Índice de Commodities - Brasil (IC-Br)	BCB
Custo Médio Construção Civil	BCB
Pesquisa Mensal do Comércio (PMC)	BCB
Clima: ONI (El Nino)	RSOI
Brasil: Taxa Nominal 3M	B3
Brasil: Taxa Nominal 1Y	B3
Brasil: Taxa Nominal 2Y	B3
Brasil: Taxa Nominal 5Y	B3
Brasil: Taxa FRA 1Y1Y	B3
Brasil: Taxa Real 1Y	B3
Brasil: Taxa Real 2Y	B3
Brasil: Taxa Real 5Y	B3
Brasil: Inflação Implícita 1Y	autor
Brasil: Inflação Implícita 2Y	autor
Brasil: Inflação Implícita 5Y	autor
Brasil: FRA Inflação Implícita 1Y2Y	autor
Brasil: Meta Inflação 1.5Y Interpolada	autor

Brasil: FRA Taxa Nominal 1Y1Y - Taxa Nominal 3M	autor
Brasil: FRA Inflação Implícita 1Y2Y - Meta Inflação 1.5Y Interpolada	autor
Brasil: CDS 5Y	Bloomberg
Ibovespa	Bloomberg
CITI BRL REER	Bloomberg
Brasil Financial Conditions	Bloomberg
Brasil CITI Terms of Trade	Bloomberg
Brasil CITI CESI (Economic Surprise Index)	Bloomberg
USD: Taxa Nominal 1Y	Bloomberg
USD: Taxa Nominal 2Y	Bloomberg
USD: Taxa Nominal 5Y	Bloomberg
USD: Inflação Implícita 1Y	Bloomberg
USD: Inflação Implícita 2Y	Bloomberg
USD: Inflação Implícita 5Y	Bloomberg
DXY	Bloomberg
US Corporate BAA 10Y Spread	Bloomberg
Gold	Bloomberg
Baltic Dry	Bloomberg
S&P500	Bloomberg
Nasdaq	Bloomberg
Eurostoxx	Bloomberg

Fonte: autor

4. MODELOS

Nesta seção, vamos fornecer mais informações sobre os modelos que foram utilizados para realizar as previsões.

4.1 BENCHMARK

O modelo ARIMA é um modelo estatístico utilizado para previsão de séries temporais. Ele é uma combinação de um modelo auto-regressivo (AR), um modelo de média móvel (MA) e uma diferença integrada (I). O modelo AR considera que o valor atual da série temporal depende das observações anteriores da mesma série, enquanto o modelo MA leva em conta a média dos resíduos dos termos anteriores. Já a diferença integrada representa a quantidade de diferenciação necessária para tornar a série temporal estacionária.

O modelo ARIMA é amplamente utilizado como um modelo de benchmark para previsão de séries temporais, especialmente quando se deseja comparar a performance de modelos mais complexos. É um modelo relativamente simples que pode ser facilmente implementado e interpretado, mas sua capacidade de previsão pode ser limitada em comparação com modelos mais sofisticados.

4.2 LINEAR

O modelo linear é um modelo de regressão que assume uma relação linear entre a variável dependente e as variáveis explicativas. O modelo linear pode ser utilizado para previsões de curto prazo e para análises de sensibilidade, pois permite identificar a relação entre a variável dependente e as variáveis explicativas. No entanto, o modelo linear pode ser limitado em termos de precisão de previsão em séries temporais complexas como a inflação, pois assume uma relação linear entre as variáveis.

A regressão linear é o mais simples dos modelos utlizados nestes estudo. Este modelo encontra os parâmetros para β que minimizam o EQM (Erro Quadrático Médio). O EQM é calculato através da soma das diferenças ao quadrado entre a previsão e os valores reais.

Também é importante destacar que o modelo linear pode não ser capaz de capturar relações não-lineares entre as variáveis, o que pode limitar sua precisão em séries temporais mais complexas. Além disso, o modelo linear pode ser sensível a outliers e a mudanças abruptas nas séries temporais, o que pode impactar negativamente a qualidade das previsões. Por isso, é importante considerar modelos mais sofisticados, como os modelos Regularização, Árvores de Decisão e Máquinas de Vetor de Suporte.

4.3 REGULARIZAÇÃO

Os métodos de regularização são utilizados para evitar o overfitting em modelos de regressão, reduzindo a magnitude dos coeficientes e melhorando a generalização do modelo. Eles partem do problema de otimização dos mínimos quadrados, adicionando um termo que penaliza a magnitude dos coeficientes. O objetivo é encontrar o vetor de coeficientes preditores β que minimiza a soma dos resíduos ao quadrado e o termo de penalização. Enquanto o hiperparâmetro λ controla a penalidade da equação, o hiperparâmetro θ define a o método de regularização.

$$\min_{\beta} \frac{1}{T}\|\boldsymbol{\pi}_n - \boldsymbol{\beta}'\boldsymbol{X}\|_2^2 + \lambda\left(\theta\|\boldsymbol{\beta}\|_1 + \frac{1}{2}(1-\theta)\|\boldsymbol{\beta}\|_2^2\right) \tag{1}$$

O método **Lasso** é uma técnica de regularização que introduz um termo de penalização L1 na função de custo da regressão linear, o que faz com que a solução resultante tenha muitos coeficientes iguais a zero. Isso torna o modelo mais simples, selecionando apenas as variáveis mais relevantes para a previsão. Além disso, o Lasso é particularmente útil em problemas de alta dimensionalidade, em que há muitas variáveis explicativas, mas apenas algumas delas são relevantes para a previsão. Essa técnica pode ajudar a melhorar a precisão do modelo, evitando o overfitting e aumentando a sua capacidade de generalização.

O método **Ridge** é outro método de regularização que adiciona um termo de penalização à função objetivo de mínimos quadrados. Esse termo de penalização é definido como a soma do quadrado dos valores dos coeficientes multiplicados por um hiperparâmetro λ. Assim como o Lasso, o método Ridge também reduz a magnitude dos coeficientes, mas de forma mais suave, sem forçar que alguns coeficientes sejam exatamente zero. Esse método é especialmente útil quando há multicolinearidade entre as variáveis explicativas, ou seja, quando elas estão altamente correlacionadas. Nesses casos, o método Ridge pode ajudar a reduzir a variância dos coeficientes, melhorando a estabilidade das estimativas. Uma vantagem adicional do método Ridge é que ele pode ser facilmente aplicado a problemas de regressão linear com mais variáveis explicativas do que observações, que são comuns em problemas de aprendizado de máquina.

Já o método **Elastic Net** é uma combinação dos métodos Ridge e Lasso, que combina as suas vantagens. Ele introduz dois hiperparâmetros, um para controlar a magnitude da penalização L1 e outro para controlar a magnitude da penalização L2. O objetivo é encontrar um equilíbrio entre a seleção de variáveis importantes e a estabilização das estimativas dos coeficientes.

4.4 ÁRVORES DE DECISÃO

Uma árvore de decisão é um modelo preditivo que usa uma estrutura em forma de árvore para representar as decisões e suas possíveis consequências. Cada nó da árvore representa uma decisão ou um teste em uma das variáveis explicativas, e as arestas representam as possíveis respostas ou saídas. As folhas da árvore representam as decisões finais ou previsões.

Na construção da árvore de decisão, o algoritmo começa com o conjunto completo de dados e tenta encontrar a variável que melhor separa as observações em grupos distintos. A separação é medida por uma métrica, como a impureza de Gini ou a entropia, e o algoritmo tenta minimizá-la. Em seguida, o algoritmo aplica o mesmo procedimento para cada grupo recém-criado, dividindo-o em subgrupos menores. O processo continua até que a árvore esteja completa ou atenda a algum critério de parada, como a profundidade máxima da árvore ou o número mínimo de observações em cada folha.

As árvores de decisão têm a vantagem de serem facilmente interpretáveis e explicáveis, pois as decisões e suas razões são apresentadas de forma clara e intuitiva. Além disso, as árvores de decisão podem lidar com variáveis categóricas e numéricas, e não requerem muita preparação dos dados.

No entanto, as árvores de decisão também têm algumas desvantagens, como a tendência ao overfitting, especialmente em árvores profundas ou em conjuntos de dados pequenos. Para mitigar esse problema, podem ser usadas técnicas de poda ou de regularização, como o método de bagging ou o método de boosting. Outra maneira de melhorar a eficácia da árvore de decisão é através da combinação de múltiplas árvores de decisão, como no caso do modelo Random Forest, LightGBM e XGBoost.

4.4.1 RANDOM FOREST

O modelo Random Forest (RF) foi introduzido por Breiman (2001) e consiste em uma combinação de várias árvores de decisão para fazer previsões. Essas árvores são construídas utilizando amostragem aleatória com reposição (bootstrap) e subconjuntos aleatórios de variáveis (features) em cada nó da árvore, o que introduz aleatoriedade e reduz o overfitting.

O processo de treinamento do modelo Random Forest começa com a construção de várias árvores de decisão independentes e, em seguida, combina as previsões dessas árvores para gerar uma previsão final. No caso de problemas de regressão, a previsão final é a média das previsões das árvores; no caso de problemas de classificação, a previsão final é a classe

mais comum entre as previsões das árvores. Uma das principais vantagens do Random Forest é que ele é capaz de lidar com conjuntos de dados muito grandes e com muitas variáveis (features), além de ser relativamente robusto a dados ausentes e outliers. Além disso, o modelo não requer muita preparação dos dados, como normalização ou tratamento de valores ausentes.

No entanto, o Random Forest também tem algumas desvantagens. O modelo pode ser computacionalmente intensivo, especialmente para grandes conjuntos de dados e número de árvores. Além disso, a interpretação do modelo pode ser mais difícil em comparação com modelos mais simples, como a regressão linear ou a árvore de decisão única.

4.4.2 LIGHT GBM

O LightGBM, introduzido por Ke et al. (2017), é um modelo que cresce verticalmente escolhendo, folha a folha, aquela que apresenta o máximo delta loss para crescer. Ele é considerado uma versão otimizada do algoritmo Gradient Boosting Machine (GBM) e usa a técnica de gradient boosting para melhorar o desempenho do modelo. A principal diferença entre Light GBM e GBM é que o Light GBM usa um algoritmo de divisão de folha em vez de divisão de nó para construir as árvores.

O Light GBM é especialmente adequado para conjuntos de dados grandes e esparsos, pois é capaz de lidar com esses conjuntos de dados com mais eficiência do que outros algoritmos de árvore de decisão. Isso se deve ao fato de que ele usa uma técnica de amostragem de dados chamada Gradient-based One-Side Sampling (GOSS) para selecionar amostras importantes e reduzir a complexidade do modelo. Além disso, o Light GBM pode usar GPUs para acelerar o processo de treinamento.

Uma das vantagens do Light GBM é que ele produz resultados de alta precisão em uma variedade de conjuntos de dados. Ele também é facilmente personalizável, permitindo que os usuários ajustem vários parâmetros, como taxa de aprendizado, profundidade da árvore e número de amostras, para otimizar o desempenho do modelo. No entanto, como outros algoritmos de árvore de decisão, o Light GBM pode sofrer de overfitting se não for adequadamente regularizado.

4.4.3 XGBOOST

XGBoost é um modelo de aprendizado de máquina baseado em árvore de decisão introduzido por Chen e Guestrin (2016). É uma implementação otimizada do algoritmo Gradient Boosting Machine (GBM) que usa uma abordagem de boosting para melhorar o desempenho do modelo. O XGBoost é conhecido por ser um dos modelos mais poderosos e populares em competições de ciência de dados.

O XGBoost usa uma técnica de gradient boosting que ajusta sucessivamente os resíduos da árvore anterior para construir uma nova árvore. Isso significa que cada nova árvore aprende a prever os erros deixados pelas árvores anteriores. O XGBoost usa uma função objetivo regularizada para evitar o overfitting e pode ser usado para problemas de classificação, regressão e ranking.

Uma das principais vantagens do XGBoost é que ele é altamente escalável e pode lidar com conjuntos de dados grandes com muitas variáveis. Além disso, ele usa uma técnica de amostragem de linha e de recurso para reduzir o overfitting e aumentar a generalização. O XGBoost também é conhecido por produzir resultados de alta precisão em competições de ciência de dados e é facilmente personalizável, permitindo que os usuários ajustem vários parâmetros para otimizar o desempenho do modelo.

No entanto, o XGBoost também tem algumas desvantagens, como o fato de que pode ser computacionalmente intensivo e exigir recursos de hardware avançados, como GPUs. Além disso, pode ser mais difícil de interpretar em comparação com modelos mais simples, como a regressão linear ou a árvore de decisão única.

4.5 MÁQUINAS DE VETOR DE SUPORTE (SVM)

As SVMs foram introduzidas por Vapnik e Chervonenkis (1963) e são um modelo de aprendizado de máquina supervisionado utilizado para classificação e regressão. O objetivo das SVM é encontrar um hiperplano que separe os dados em duas classes de forma ótima. A ideia principal das SVM é mapear os dados de entrada em um espaço de alta dimensão e encontrar o hiperplano que melhor separa as duas classes, maximizando a margem entre os pontos de cada classe e o hiperplano. Essa margem é definida como a distância perpendicular entre o hiperplano e o ponto mais próximo de cada classe, e as SVMs buscam encontrar o hiperplano que maximize essa distância.

Uma das vantagens das SVMs é que elas podem lidar com conjuntos de dados de alta dimensão e são capazes de produzir resultados de alta precisão em muitos problemas de classificação e regressão. As SVMs também podem ser usadas com diferentes funções de kernel para lidar com dados não lineares, como a função de kernel RBF (radial basis function) que permite lidar com dados não lineares mapeando-os em um espaço de alta dimensão.

No entanto, as SVMs podem ser computacionalmente intensivas, especialmente quando o conjunto de dados é grande. Além disso, as SVMs podem ser sensíveis à escolha de hiperparâmetros e à qualidade dos dados, o que pode afetar o desempenho do modelo. Por fim, as SVMs podem ser difíceis de interpretar e entender em comparação com outros modelos mais simples, como árvores de decisão.

O algoritmo SVM tem como objetivo encontrar um hiperplano em um espaço N-dimensional que separe distintamente os pontos de dados. Existem duas principais variantes do SVM: SVM Linear e SVM não-linear.

4.5.1 SVM LINEAR

As SVMs lineares são um tipo de SVM em que o hiperplano de separação é uma linha reta. Essa linha reta é definida como a combinação linear dos vetores de suporte, que são os pontos mais próximos das duas classes. O objetivo das SVMs lineares é encontrar a linha que separa as duas classes com a maior margem possível, ou seja, a maior distância entre a linha e os pontos de cada classe.

Uma das principais vantagens das SVMs lineares é a sua eficiência computacional e escalabilidade para grandes conjuntos de dados. Além disso, por ter uma abordagem simples, é fácil de interpretar e entender os resultados do modelo. No entanto, as SVMs lineares podem ter um desempenho inferior em dados não linearmente separáveis, onde é necessário utilizar outros tipos de SVMs, como as SVMs com kernel.

4.5.1 KERNEL / NÃO-LINEAR

Dados não lineares que não permitem separação em categorias distintas através de um vetor linear são classificados usando um kernel ou SVM não-linear. Aqui, o classificador é chamado de classificador não linear. A classificação é feita com um tipo de dados não linear e agregando dimensões maiores. Alguns parâmetos do SVM são fundamentais para esta

aplicação: o parâmetro C é responsável por controlar o nível de tolerância aos erros do modelo. Já o parâmetro Gamma está presente somente no SVM não linear, através das funções de Kernel (especificamente o Radial, Polinomial e Sigmoide). Neste estudo, vamos nos limitar apenas aos Kernel Radial e Polinomial.

5. RESULTADOS

Nesta seção, discutiremos os resultados de cada desagregação proposta neste estudo e apresentaremos as respectivas conclusões.

5.1 IPCA

Tabela 3 - MAE (Mean Absolute Error) - IPCA

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.2027	0.2034	0.2023	0.2050	0.2075	0.2047
1	LINEAR	0.2642	0.2405	0.2829	0.3711	0.3549	0.3530
2	LASSO	0.3554	0.2781	0.3415	0.3250	0.3120	0.3162
3	RIDGE	0.2444	0.1782	0.2067	0.1997	0.1874	0.1836
4	ELASTICNET	0.1793	0.1733	0.1596	0.1391	0.1371	0.1353
5	RANDOM FOREST	0.1378	0.0855	0.0972	0.0731	0.0674	0.0663
6	SVM LINEAR	0.0442	0.1206	0.1482	0.1349	0.1464	0.1432
7	SVM POLINOMIAL	0.2429	0.1768	0.2162	0.2149	0.1973	0.2061
8	SVM RADIAL	0.2535	0.1567	0.1925	0.1727	0.1658	0.1770
9	LightGBM	0.2757	0.2087	0.2342	0.2107	0.1991	0.2049
10	XGBoost	0.1853	0.1522	0.1570	0.1439	0.1341	0.1449

No horizonte de 1 mês, o modelo SVM Linear obteve o menor erro absoluto médio (MAE), o que o coloca como o melhor modelo para esse período específico. Em segundo lugar, o Random Forest teve um desempenho relativamente bom, embora não tenha superado o desempenho do SVM Linear.

No entanto, ao considerar todos os outros horizontes de previsão observados, o modelo Random Forest se destacou como o melhor modelo em termos de MAE. Isso indica que, para horizontes de previsão mais longos, o Random Forest apresentou um desempenho geral superior em comparação com todos os outros modelos estudados.

A diferença de desempenho entre o SVM Linear e o Random Forest nos horizontes de previsão pode ser atribuída às características de cada modelo. O SVM Linear é um modelo de aprendizado de máquina que busca encontrar um hiperplano que separe os dados de maneira

ótima. Ele é particularmente adequado para problemas com separação linear e pode ser eficaz quando as relações entre as variáveis de entrada e a variável de saída são lineares. No caso do horizonte de 1 mês, o SVM Linear foi capaz de capturar essas relações lineares de maneira mais precisa, resultando em um menor erro absoluto médio. Portanto, nos horizontes de previsão mais longos, nos quais as relações podem ser mais complexas e não lineares, o Random Forest pode ter uma vantagem sobre o SVM Linear.

5.2 ADMINISTRADOS

Tabela 4 - MAE (Mean Absolute Error) - Administrados

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.4636	0.4692	0.4724	0.4774	0.4792	0.4763
1	LINEAR	0.0257	0.4809	0.9467	1.7213	1.5686	1.6550
2	LASSO	0.7546	0.6149	0.7540	0.7007	0.6767	0.7205
3	RIDGE	0.6957	0.5489	0.6352	0.5569	0.5420	0.5595
4	ELASTICNET	0.3613	0.3773	0.4002	0.3251	0.3283	0.3312
5	RANDOM FOREST	0.3104	0.2454	0.3403	0.2119	0.2098	0.2109
6	SVM LINEAR	0.1904	0.3164	0.3075	0.3197	0.3233	0.3268
7	SVM POLINOMIAL	0.7467	0.5804	0.7126	0.6548	0.6259	0.6304
8	SVM RADIAL	0.6184	0.4430	0.5767	0.5165	0.4839	0.5357
9	LightGBM	0.7410	0.5701	0.6589	0.5701	0.5531	0.5716
10	XGBoost	0.5958	0.5226	0.5302	0.3865	0.3850	0.4139

No horizonte de 1 mês, o modelo Linear obteve o menor erro absoluto médio (MAE), o que o coloca como o melhor modelo para esse período específico. Isso significa que o modelo Linear foi capaz de fazer previsões mais precisas em comparação com os outros modelos avaliados.

No entanto, ao considerar todos os outros horizontes de previsão observados, os modelos Random Forest e SVM Linear se destacaram como os melhores modelos em termos de MAE. Esses dois modelos alternaram entre si a posição de destaque em todos os outros horizontes de previsão, o que indica que eles foram consistentemente melhores na previsão de dados para diferentes períodos.

5.3 LIVRES

Tabela 5 - MAE (Mean Absolute Error) – Livres

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.2088	0.2082	0.2072	0.2097	0.2111	0.2083
1	LINEAR	0.2752	0.2287	0.3063	0.4265	0.4325	0.4275
2	LASSO	0.3152	0.2468	0.2843	0.2995	0.3087	0.2999
3	RIDGE	0.2344	0.1733	0.1806	0.1921	0.1868	0.1915
4	ELASTICNET	0.1112	0.1084	0.1082	0.1007	0.1012	0.1014
5	RANDOM FOREST	0.0825	0.0644	0.0452	0.0541	0.0576	0.0559
6	SVM LINEAR	0.0377	0.0900	0.1526	0.1062	0.1184	0.1157
7	SVM POLINOMIAL	0.2335	0.1685	0.1810	0.1997	0.1918	0.1978
8	SVM RADIAL	0.2237	0.1598	0.1402	0.1475	0.1560	0.1497
9	LightGBM	0.2068	0.1689	0.1652	0.1848	0.1830	0.1859
10	XGBoost	0.1104	0.1314	0.1114	0.1155	0.1194	0.1185

No horizonte de 1 mês, o modelo SVM Linear obteve o menor erro absoluto médio (MAE), o que o coloca como o melhor modelo para esse período específico. Em segundo lugar, o Random Forest teve um desempenho relativamente bom, embora não tenha superado o desempenho do SVM Linear.

No horizonte de 3 meses, o melhor modelo foi o Random Forest, seguido pelo SVM Linear em segundo lugar. No entanto, ao considerar todos os outros horizontes de previsão observados, o modelo Random Forest se destacou como o melhor modelo em termos de MAE, seguido pelo Elastic Net. Isso indica que, para horizontes de previsão mais longos, o Random Forest apresentou um desempenho geral superior em comparação com todos os outros modelos estudados.

5.4 ALIMENTAÇÃO NO DOMICÍLIO

Tabela 6 - MAE (Mean Absolute Error) - Alimentação no Domicílio

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.2042	0.2016	0.2019	0.2035	0.2030	0.2038
1	LINEAR	0.0537	0.1270	0.2714	0.3296	0.3316	0.3276
2	LASSO	0.3300	0.3110	0.3238	0.3054	0.3081	0.3024
3	RIDGE	0.2306	0.1833	0.1916	0.1889	0.1894	0.1812
4	ELASTICNET	0.1621	0.0979	0.1103	0.1055	0.1061	0.1093
5	RANDOM FOREST	0.0898	0.0674	0.0712	0.1030	0.1098	0.1014
6	SVM LINEAR	0.1579	0.0933	0.0945	0.0837	0.0866	0.0863
7	SVM POLINOMIAL	0.2412	0.1832	0.2003	0.2005	0.1923	0.1954
8	SVM RADIAL	0.2250	0.1917	0.1916	0.1715	0.1767	0.1908
9	LightGBM	0.2194	0.2067	0.2038	0.2307	0.2375	0.2205
10	XGBoost	0.1320	0.1432	0.1337	0.1385	0.1491	0.1394

No horizonte de 1 mês, o modelo Linear obteve o menor erro absoluto médio (MAE), o que o coloca como o melhor modelo para esse período específico. Isso significa que o modelo Linear foi capaz de fazer previsões mais precisas em comparação com os outros modelos avaliados. Em segundo lugar, o Random Forest também teve um bom desempenho, seguido pelo XGBoost em terceiro lugar.

No entanto, ao considerar todos os outros horizontes de previsão observados, os modelos Random Forest e SVM Linear se destacaram como os melhores modelos em termos de MAE. Esses dois modelos alternaram entre si a posição de destaque em todos os outros horizontes de previsão, o que indica que eles foram consistentemente melhores na previsão de dados para diferentes períodos.

5.5 SERVIÇOS

Tabela 7 - MAE (Mean Absolute Error) - Serviços

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.2421	0.2454	0.2479	0.2502	0.2470	0.2478
1	LINEAR	0.4870	1.1025	1.2925	0.6871	0.6687	0.6779
2	LASSO	0.1395	0.1942	0.2888	0.3077	0.2971	0.2936
3	RIDGE	0.1442	0.1757	0.1934	0.2074	0.2049	0.2062
4	ELASTICNET	0.0734	0.0799	0.1072	0.1128	0.1143	0.1113
5	RANDOM FOREST	0.0315	0.0506	0.0569	0.0788	0.0710	0.0779
6	SVM LINEAR	0.2233	0.1042	0.1235	0.1376	0.1369	0.1358
7	SVM POLINOMIAL	0.1076	0.1553	0.1969	0.2140	0.2146	0.2153
8	SVM RADIAL	0.0648	0.1201	0.1518	0.1540	0.1440	0.1430
9	LightGBM	0.1470	0.1841	0.2053	0.2140	0.2025	0.2083
10	XGBoost	0.0588	0.1170	0.1242	0.1433	0.1376	0.1482

Em todos os horizontes de previsão observados, o modelo Random Forest se destacou como o melhor em termos de MAE. Sua superioridade pode ser constatada em todos os horizontes estudados, demonstrando sua capacidade preditiva consistente para o setor de Serviços.

Além disso, é importante ressaltar a capacidade preditiva do modelo Elastic Net. Este modelo ficou em segundo lugar em praticamente todos os prazos observados, o que indica que também possui um bom desempenho na previsão de dados para diferentes períodos.

5.6 INDUSTRIAIS

Tabela 8 - MAE (Mean Absolute Error) - Industriais

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.6514	0.6523	0.6519	0.6557	0.6585	0.6678
1	LINEAR	1.8616	2.6286	2.6765	1.1700	1.2551	1.2326
2	LASSO	0.9504	0.7356	0.7415	0.7575	0.8085	0.7992
3	RIDGE	0.5832	0.5558	0.5129	0.5191	0.4932	0.5162
4	ELASTICNET	0.4115	0.4846	0.4049	0.2807	0.2882	0.3041
5	RANDOM FOREST	0.2715	0.2195	0.1648	0.1385	0.1440	0.1423
6	SVM LINEAR	0.2083	0.3658	0.6210	0.3122	0.3576	0.3303
7	SVM POLINOMIAL	0.6292	0.5762	0.5444	0.5612	0.5429	0.5529
8	SVM RADIAL	0.6085	0.4515	0.3750	0.4104	0.4296	0.4152
9	LightGBM	0.6346	0.5348	0.5139	0.5266	0.5218	0.5247
10	XGBoost	0.4400	0.4427	0.4018	0.4279	0.4130	0.4173

No horizonte de 1 mês, o modelo SVM Linear obteve o menor erro absoluto médio (MAE), o que o coloca como o melhor modelo para esse período específico. Em segundo lugar, o Random Forest também teve um bom desempenho, seguido pelo Elastic Net em terceiro lugar.

No entanto, ao considerar todos os outros horizontes de previsão observados, o modelo Random Forest se destacou como o melhor modelo em termos de MAE. Em muitos desses períodos, os modelos Elastic Net e SVM Linear também tiveram um bom desempenho, mostrando que eles são alternativas competitivas para a previsão de dados em diferentes períodos.

5.7 TRADABLES

Tabela 9 - MAE (Mean Absolute Error) - Tradables

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.2726	0.2725	0.2727	0.2755	0.2761	0.2758
1	LINEAR	0.8351	0.6659	0.8735	0.8824	0.7965	0.8599
2	LASSO	0.4414	0.3834	0.3598	0.3690	0.4021	0.3970
3	RIDGE	0.2960	0.1969	0.1895	0.1994	0.2210	0.2202
4	ELASTICNET	0.1884	0.1778	0.1819	0.1348	0.1391	0.1419
5	RANDOM FOREST	0.1018	0.0720	0.0652	0.0701	0.0855	0.0812
6	SVM LINEAR	0.1497	0.1714	0.2353	0.1493	0.1686	0.1748
7	SVM POLINOMIAL	0.2769	0.1927	0.1806	0.2084	0.2321	0.2232
8	SVM RADIAL	0.2949	0.1936	0.1538	0.1743	0.2090	0.1891
9	LightGBM	0.2609	0.2037	0.1961	0.2037	0.2310	0.2187
10	XGBoost	0.1674	0.1258	0.1219	0.1350	0.1612	0.1594

O modelo Random Forest demonstrou consistência e um desempenho superior em todos os horizontes observados. Essa é uma característica muito desejável em um modelo de previsão, pois indica sua capacidade de generalizar bem para diferentes períodos de tempo.

O modelo XGBoost também apresentou bons resultados nos horizontes de até 6 meses. O XGBoost é conhecido por sua capacidade de lidar com conjuntos de dados complexos e fornecer previsões precisas. Seu desempenho em horizontes mais curtos indica que ele é capaz de capturar padrões e tendências de curto prazo nos dados.

Por outro lado, o Elastic Net se destacou nos horizontes mais longos, após 6 meses. O Elastic Net é um modelo de regressão linear regularizado que combina as propriedades do Lasso e do Ridge. Sua capacidade de lidar com conjuntos de dados com multicolinearidade e de selecionar automaticamente variáveis relevantes pode ser especialmente útil em previsões de longo prazo, onde os padrões podem ser mais complexos.

5.8 NON-TRADABLES

Tabela 10 - MAE (Mean Absolute Error) - Non-Tradables

		h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
0	BENCHMARK (ARIMA)	0.2655	0.2662	0.2666	0.2678	0.2653	0.2686
1	LINEAR	0.2264	0.3318	0.5505	0.8385	0.7593	0.7789
2	LASSO	0.3077	0.2521	0.3431	0.3550	0.3410	0.3494
3	RIDGE	0.2186	0.1832	0.2216	0.2322	0.2221	0.2372
4	ELASTICNET	0.0989	0.0923	0.1067	0.1191	0.1212	0.1257
5	RANDOM FOREST	0.0904	0.0697	0.0731	0.0832	0.0744	0.0828
6	SVM LINEAR	0.1187	0.0789	0.1136	0.1326	0.1296	0.1313
7	SVM POLINOMIAL	0.2207	0.1805	0.2381	0.2491	0.2324	0.2438
8	SVM RADIAL	0.1848	0.1496	0.1857	0.1917	0.1750	0.1866
9	LightGBM	0.2169	0.1878	0.2412	0.2524	0.2319	0.2444
10	XGBoost	0.1545	0.1338	0.1614	0.1650	0.1561	0.1589

O modelo Random Forest demonstrou consistência e um desempenho superior em todos os horizontes observados. Essa característica é muito desejável em um modelo de previsão, pois indica sua capacidade de generalizar bem para diferentes períodos de tempo.

Além do Random Forest, o desempenho do Elastic Net também foi destacado, ficando em segundo lugar na maioria dos horizontes observados. Sua capacidade de lidar com multicolinearidade nos dados e selecionar automaticamente as variáveis mais relevantes é uma vantagem importante para previsões precisas.

Em terceiro lugar, o modelo SVM Linear também mostrou um bom desempenho robusto e consistente em praticamente todos os períodos estudados. Sua capacidade de lidar com conjuntos de dados de alta dimensionalidade e sua interpretabilidade são vantagens importantes.

5.9 DESAGREGAÇÕES

Tabela 11 - MAE (Mean Absolute Error) - Desagregações - ARIMA

	h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
AGG 1	0.2027	0.2034	0.2023	0.2050	0.2075	0.2079
AGG 2	0.2006	0.2017	0.2008	0.2032	0.2057	0.2049
AGG 3	0.2064	0.2075	0.2064	0.2087	0.2113	0.2099
AGG 4	0.1971	0.1980	0.1971	0.1986	0.2015	0.2017

A tabela 11 apresenta os valores do erro absoluto médio (MAE) para diferentes horizontes de previsão utilizando o modelo ARIMA e quatro agregações distintas (AGG 1, AGG 2, AGG 3 e AGG 4). Cada valor na tabela corresponde ao MAE para um determinado horizonte de previsão, sendo eles: 1, 3, 6, 9, 12 e 18 meses.

Ao analisar os resultados da tabela, podemos observar que o modelo AGG 4 obteve o melhor desempenho em todos os horizontes de previsão estudados. Em todos os prazos, o AGG 4 apresentou os menores valores de MAE, indicando sua maior precisão nas previsões em comparação com os outros modelos.

O modelo AGG 2 também teve um desempenho consideravelmente bom, com valores de MAE próximos aos do AGG 4 em todos os horizontes observados. Isso demonstra que o AGG 2 foi capaz de fazer previsões precisas em diferentes períodos de tempo.

O AGG 1, por sua vez, também apresentou resultados satisfatórios, embora com MAEs ligeiramente maiores que os dos modelos AGG 4 e AGG 2. Ainda assim, o AGG 1 demonstrou ser um modelo competente na previsão dos dados para os horizontes considerados.

Por outro lado, o modelo AGG 3 mostrou-se menos eficaz em todos os horizontes de previsão analisados, com MAEs mais altos em comparação com os outros modelos. Isso indica que o AGG 3 teve um desempenho inferior na previsão dos dados para esses períodos específicos.

Portanto, com base nos resultados da tabela, concluímos que o modelo AGG 4 foi o mais preciso em termos de MAE, seguido pelo AGG 2 e AGG 1, enquanto o modelo AGG 3 teve o desempenho menos satisfatório entre os modelos avaliados.

Tabela 12 - MAE (Mean Absolute Error) - Desagregações MACHINE LEARNING

	h = 1	h = 3	h = 6	h = 9	h = 12	h = 18
AGG 1	0.0600	0.0344	0.0344	0.0556	0.0556	0.0558
AGG 2	0.0576	0.0403	0.0402	0.0597	0.0597	0.0598
AGG 3	0.1088	0.0885	0.0818	0.0977	0.0977	0.0987
AGG 4	0.0421	0.0419	0.0421	0.0622	0.0622	0.0625

A tabela 12 apresenta os valores do erro absoluto médio (MAE) para diferentes horizontes de previsão utilizando o modelo MACHINE LEARNING e quatro agregações distintas (AGG 1, AGG 2, AGG 3 e AGG 4). Cada valor na tabela corresponde ao MAE para um determinado horizonte de previsão, sendo eles: 1, 3, 6, 9, 12 e 18 meses.

Ao analisar os resultados da tabela 12, podemos observar que o modelo AGG 4 obteve o menor erro absoluto médio (MAE) quando o horizonte de previsão é de 1 mês, indicando seu desempenho superior em comparação com os outros modelos nesse período específico. Em segundo lugar, o modelo AGG 2 apresentou um MAE ligeiramente maior que o AGG 4, porém ainda demonstrando um bom desempenho. Em terceiro lugar, o modelo AGG 1 mostrou resultados satisfatórios, embora com um MAE um pouco maior que os modelos AGG 4 e AGG 2.

No entanto, quando o horizonte de previsão passa a ser a partir de 3 meses, os resultados se modificam. O modelo AGG 1 se destaca como o mais consistente, apresentando o menor MAE em todos os horizontes de previsão a partir de 3 meses. Em segundo lugar, temos o modelo AGG 2, seguido pelo AGG 4.

Esses resultados sugerem que o modelo AGG 1 é mais eficaz em prever os dados quando o horizonte de previsão é a partir de 3 meses, mostrando consistência em seu desempenho. O modelo AGG 2 também apresenta um desempenho satisfatório nesses horizontes, enquanto o AGG 3 demonstra um desempenho um pouco inferior, com MAEs mais altos em comparação com os outros modelos.

6. CONCLUSÕES

Ao analisar os resultados das tabela 11 e 12, podemos observar que a Agregação 4 (AGG 4) obteve o menor erro absoluto médio (MAE) no primeiro horizonte de previsão (h=1), tanto no benchmark quanto nos modelos de Machine Learning. Isso indica que a AGG 4 é a melhor opção de estimação nesse horizonte, apresentando uma precisão superior em comparação com as outras agregações.

É interessante notar que, para esse horizonte de previsão, os modelos de Machine Learning superaram o benchmark em todas as agregações estudadas. Isso sugere que há um ganho significativo ao utilizar os modelos de Machine Learning, especialmente quando combinados com a AGG 4.

No entanto, para os demais horizontes de previsão (h=3, 6, 9 e 12), não há ganhos significativos ao utilizar as diferentes agregações nos modelos de Machine Learning. Apesar disso, os modelos de Machine Learning ainda apresentam resultados consideravelmente melhores que o benchmark, mesmo quando utilizando a Agregação 1 (headline).

Portanto, podemos concluir que, no primeiro horizonte de previsão, a AGG 4 é a melhor opção e os modelos de Machine Learning superam o benchmark. Nos demais horizontes, a escolha da agregação tem menos impacto nos resultados dos modelos de Machine Learning, mas esses modelos ainda se mostram superiores ao benchmark.

REFERÊNCIAS

ANDREW, BEKAERT & WEI. Do macro variables, asset markets, or surveys forecast inflation better, 2007.

AHMED, N. K. et al. An empirical comparison of machine learning models for time series forecasting. Econometric Reviews, 2010.

ARAUJO, Gustavo Silva; VICENTE, José Valentim. Estimação da Inflação Implícita de Curto Prazo. Working Papers Series 460, Central Bank of Brazil, Research Department, 2017.

ARAUJO & GAGLIANONE. Machine Learning Methods for Inflation Forecasting in Brazil: new contenders versus classical models, 2022.

ATKESON & OHANIAN. Are Phillips curves useful for forecasting inflation, 2001.

BERMINGHAM & D'AGOSTINO. Understanding and forecasting aggregate and disaggregate price dynamics, 2014.

BOARETTO & MEDEIROS. Forecasting Aggregate Inflation using Disaggregates and Machine Learning, 2022.

BREIMAN, Leo. Random Forests. Machine Learning, 2001.

CARLOS, Thiago Carlomagno; MARÇAL, Emerson Fernandes. Forecasting Brazilian Inflation by Its Aggregate and Disaggregated Data: A Test of Predictive Power by Forecast Horizon.

CERQUEIRA, TORGO & MOZETIC. Evaluating time series forecasting models: an empirical study on performance estimation methods, 2019.

CERQUEIRA, TORGO & MOZETIC. Model Selection for Time Series Forecasting: Empirical Analysis of Different Estimators, 2021.

FAUST, J.; WRIGHT, J. H. Forecasting inflation. In: ELLIOTT, G.; TIMMERMANN, A. Handbook of Economic Forecasting. 1. ed. New York: Elsevier, 2013.

FIGUEIREDO, Francisco Marcos Rodrigues. Forecasting Brazilian inflation using a large data set. Working Papers Series 228, Central Bank of Brazil, Research Department, 2010.

FLORIDO, GUSTAVO. Essays On Inflation Forecasting, 2021.

FREITAS, ARAUJO & SILVA. Breakeven Inflation Rate Estimation: an alternative approach considering indexation lag and seasonality. Working Papers Series 493, Central Bank of Brazil, Research Department, 2019.

GARCIA, Márcio G. P.; MEDEIROS, Marcelo C.; VASCONCELOS, Gabriel F. R. Real-time inflation forecasting with high-dimensional models: The case of Brazil. International Journal of Forecasting, 2017.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. The elements of statistical learning. 2. ed. California: Springer, 2008.

INOUE & KILIAN. How Useful Is Bagging in Forecasting Economic Time Series? A Case Study of U.S. Consumer Price Inflation, 2008.

JAMES, G. et al. An Introduction to Statistical Learning with Application in R. 1. ed. New York: Springer New York, 2013.

MCCRACKEN & NG. FRED-MD: a Monthly Database for Macroeconomic Research, 2016.

MEDEIROS, GARCIA & VASCONCELOS. Real-time inflation forecasting with high-dimensional models: The case of Brazil, 2017.

MEDEIROS, VASCONCELOS, VEIGA, ZILBERMAN. Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods, 2019.

MONTEFORTE & MORETTI. Real Time Forecasts of Inflation: The Role of Financial Variables, 2013.

STOCK & WATSON. Forecasting Inflation, 1999.

THEOHARIDIS. Forecasting Inflation Using Deep Learning: An Application of Convolutional LSTM Networks and Variational Autoencoders, 2021.

WILLMOTT & MATSUURA. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, 2005.