



# **MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA DE SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS DE INSTITUIÇÕES FINANCEIRAS**

DEFESA DE DISSERTAÇÃO DE MESTRADO

Alberson Miranda, PPGeco/UFES

Fevereiro de 2024



# SUMÁRIO

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

### ► CONTEXTUALIZAÇÃO DA PESQUISA

### ► OBJETIVOS

### ► RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

### ► METODOLOGIA

### ► RESULTADOS

### ► REFERÊNCIAS



# ANÁLISE DE SÉRIES TEMPORAIS EM ECONOMIA BANCÁRIA

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

- Problemas populares na literatura: risco de crédito, detecção de anomalias (fraude), preços de ativos, alocação ótima de numerários (SEZER; GUDELEK; OZBAYOGLU, 2019)
- Pouco sobre projeção de saldos e rendas (*guidance, budgeting*)
- Previsões equivocadas = percepção de incompetência, redução de reputação e receitas (BECCALLI et al., 2015)
- Orçamentação em empresas de muitas filiais: alta dimensionalidade e dúvidas



# ANÁLISE DE SÉRIES TEMPORAIS EM ECONOMIA BANCÁRIA

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

- Banestes: 134 agências, 78 municípios do ES, 13 microrregiões, 4 mesorregiões
- Qual a melhor forma de realizar projeções?
  - Individualmente por agência?
  - Apenas o total e distribuir proporcionalmente?
  - As informações por mesorregião ou microrregião são importantes?
  - Como usar toda a informação?



# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

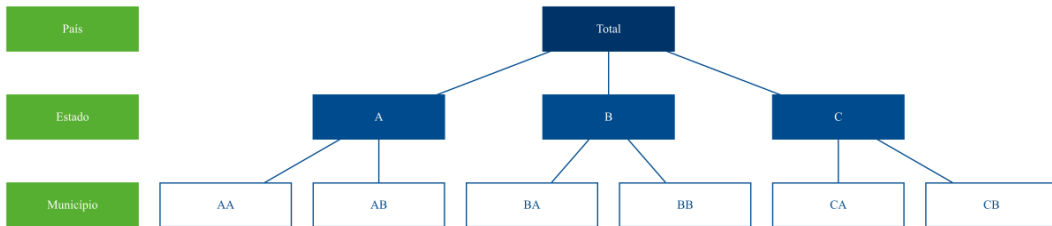
## 1 CONTEXTUALIZAÇÃO DA PESQUISA

Séries temporais hierárquicas são aquelas que podem ser agregadas ou desagregadas naturalmente em uma estrutura aninhada (HYNDMAN; ATHANASOPOULOS, 2021)



# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

## 1 CONTEXTUALIZAÇÃO DA PESQUISA





# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

- Cada nível traz informação diferente
- É razoável supor que exista covariância entre as previsões dos diferentes níveis
- Toda informação da estrutura pode ser útil



# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

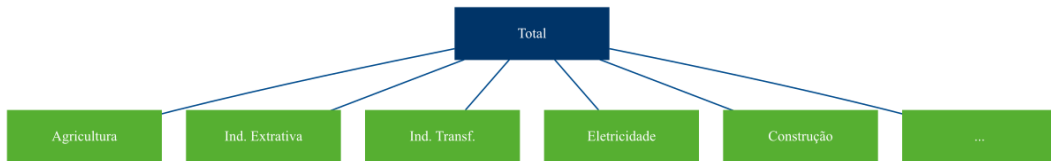
Séries temporais agrupadas são aquelas que não impõem uma única estrutura hierárquica





# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

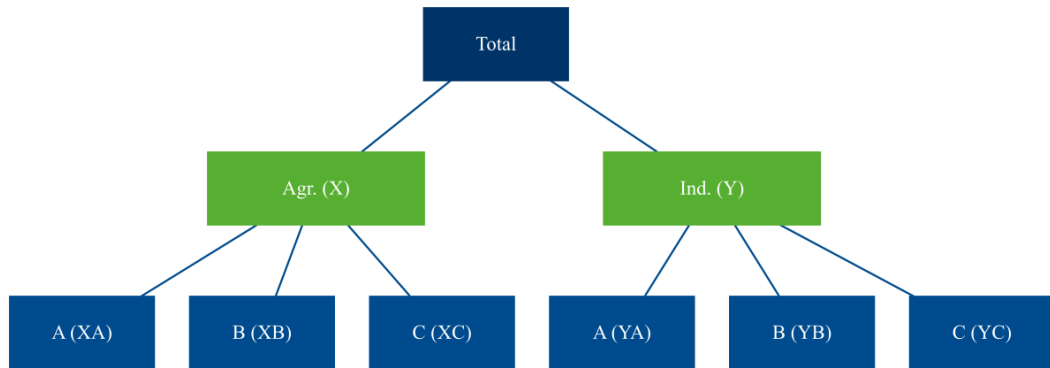
## 1 CONTEXTUALIZAÇÃO DA PESQUISA





# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

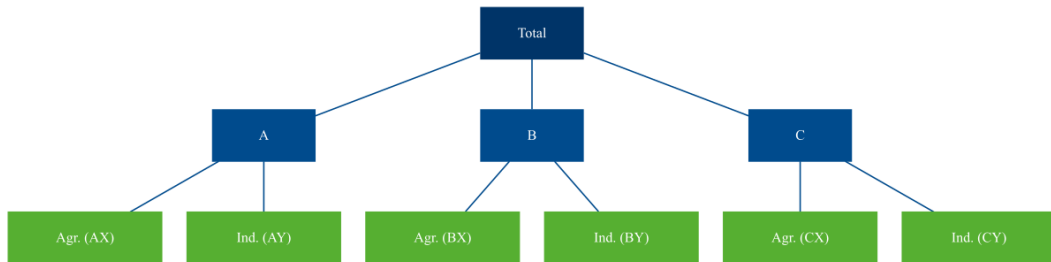
## 1 CONTEXTUALIZAÇÃO DA PESQUISA





# SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

## 1 CONTEXTUALIZAÇÃO DA PESQUISA





# COERÊNCIA E RECONCILIAÇÃO

## 1 CONTEXTUALIZAÇÃO DA PESQUISA

- Coerência: cada nó da hierarquia deve totalizar os nós filhos
- Não há razão para que as previsões individuais (e.g., Arima, ETS) sejam coerentes

### RECONCILIAÇÃO ÓTIMA

Reescrever as previsões de forma que elas sejam coerentes, utilizando toda a informação disponível na estrutura hierárquica, ao mesmo tempo em que minimiza a variância do erro de previsão



# SUMÁRIO

## 2 OBJETIVOS

► CONTEXTUALIZAÇÃO DA PESQUISA

► OBJETIVOS

► RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

► METODOLOGIA

► RESULTADOS

► REFERÊNCIAS



# OBJETIVO GERAL

## 2 OBJETIVOS

Estudar o problema da reconciliação ótima de previsões pontuais a partir de métodos de *machine learning* em séries temporais hierárquicas e agrupadas do Banco do Estado do Espírito Santo



# OBJETIVOS ESPECÍFICOS

## 2 OBJETIVOS

1. Estender a abordagem de Spiliotis et al. (2021) para a reconciliação ótima de séries temporais hierárquicas *e agrupadas*;
2. Avaliar a efetividade de mais 4 implementações de métodos de ML (*lasso*, *ridge*, *elastic net*, SVM e *lightGBM*), além dos 2 propostos no trabalho original (*XGBoost* e *random forest*);
3. Propor duas variações metodológicas para obtenção da amostra de treino dos modelos de ML;
4. Obter ganhos de performance no contexto de séries temporais financeiras.



# SUMÁRIO

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

► CONTEXTUALIZAÇÃO DA PESQUISA

► OBJETIVOS

► RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

► METODOLOGIA

► RESULTADOS

► REFERÊNCIAS





# BOTTOM-UP

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

- informações apenas dos níveis mais desagregados
- soma-se para obter os níveis agregados



# TOP-DOWN

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

- apenas informações do nível mais agregado
- há dezenas de métodos para distribuir proporcionalmente as previsões agregadas (GROSS; SOHL, 1990)
- não podem ser utilizadas no contexto de séries temporais agrupadas<sup>1</sup>.

---

<sup>1</sup>Desambiguar uma estrutura agrupada resulta em uma estrutura hierárquica, mas há necessariamente perda de informação (e.g., fixar os verbetes abaixo das agências, causaria a perda da informação do agregado dos verbetes).



# MIDDLE-OUT

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

- informações apenas de um nível intermediário
- soma-se para obter os níveis agregados e distribui-se proporcionalmente para os níveis mais desagregados



# NOTAÇÃO MATRICIAL

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

$$\tilde{y}_t = SG\hat{y}_t \quad (1)$$

Em que

- $\tilde{y}_t$  é o vetor de previsões reconciliadas
- $S$  é a matriz de soma
- $G$  é a matriz de reconciliação
- $\hat{y}_t$  é o vetor de previsões base



# MATRIZ DE SOMA

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

$S$  mapeia a estrutura hierárquica a partir da soma dos elementos mais desagregados

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \end{bmatrix}_{8 \times 1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{8 \times 5} \begin{bmatrix} \hat{y}_{AA,t} \\ \hat{y}_{AB,t} \\ \hat{y}_{AC,t} \\ \hat{y}_{BA,t} \\ \hat{y}_{BB,t} \end{bmatrix}_{5 \times 1} \quad (2)$$

*exemplo 1: matriz de soma*



# MATRIZ DE RECONCILIAÇÃO

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

$G$  mapeia o nível mais desagregado a partir das previsões de todos os níveis da hierarquia, garantindo a coerência

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \end{bmatrix}_{8 \times 1} = S_{n \times m} \begin{bmatrix} p_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{5 \times 8} \begin{bmatrix} \hat{y}_{T+h|T} \\ \hat{y}_{A,T+h|T} \\ \hat{y}_{B,T+h|T} \\ \hat{y}_{AA,T+h|T} \\ \hat{y}_{AB,T+h|T} \\ \hat{y}_{AC,T+h|T} \\ \hat{y}_{BA,T+h|T} \\ \hat{y}_{BB,T+h|T} \end{bmatrix}_{8 \times 1} \quad (3)$$

*exemplo 2: matriz de reconciliação top-down*



# MATRIZ DE RECONCILIAÇÃO

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{A,t} \\ \tilde{y}_{B,t} \\ \tilde{y}_{AA,t} \\ \tilde{y}_{AB,t} \\ \tilde{y}_{AC,t} \\ \tilde{y}_{BA,t} \\ \tilde{y}_{BB,t} \end{bmatrix}_{8 \times 1} = \mathbf{S}_{8 \times 5} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{5 \times 8} \begin{bmatrix} \hat{y}_{T+h|T} \\ \hat{y}_{A,T+h|T} \\ \hat{y}_{B,T+h|T} \\ \hat{y}_{AA,T+h|T} \\ \hat{y}_{AB,T+h|T} \\ \hat{y}_{AC,T+h|T} \\ \hat{y}_{BA,T+h|T} \\ \hat{y}_{BB,T+h|T} \end{bmatrix}_{8 \times 1} \quad (4)$$

*exemplo 3: matriz de reconciliação bottom-up*



# MATRIZ DE RECONCILIAÇÃO

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

E se quisermos utilizar toda a informação e não apenas parte do espaço das previsões base?

### O PROBLEMA DE PESQUISA DA RECONCILIAÇÃO ÓTIMA

Estimar uma matriz de reconciliação  $G$  que utilize toda a informação disponível e com o menor erro de previsão





# MATRIZ DE RECONCILIAÇÃO

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

### PROBLEMA DE REGRESSÃO

- Minimização do traço da matriz de covariância dos erros das previsões reconciliadas (MinT<sup>a</sup>)

$$\min_{\tilde{e}} y_{T+h} - \tilde{y}_{T+h} \quad (5)$$

- Mínimos Quadrados Generalizados (MQG)

$$G = (S'W_h^\dagger S)^{-1} S'W_h^\dagger \quad (6)$$

Substituindo  $G$  na Equação 1, temos:

$$\tilde{y} = S(S'W_h^\dagger S)^{-1} S'W_h^\dagger \hat{y}_h \quad (7)$$

---

<sup>a</sup>(WICKRAMASURIYA; ATHANASOPOULOS; HYNDMAN, 2019).



# RECONCILIAÇÃO POR MACHINE LEARNING

## 3 RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

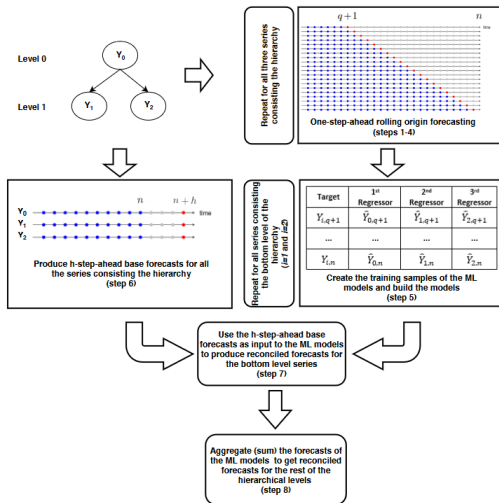


Figure 1: Abordagem de reconciliação ótima por ML. Fonte: Spiliotis et al. (2021)



# SUMÁRIO

## 4 METODOLOGIA

► CONTEXTUALIZAÇÃO DA PESQUISA

► OBJETIVOS

► RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

► **METODOLOGIA**

► RESULTADOS

► REFERÊNCIAS



# DADOS E VARIÁVEIS

## 4 METODOLOGIA

- Estatística Bancária Mensal e por Município (Bacen)
- Municípios (IBGE)
- Saldos de crédito dos verbetes empréstimos e financiamentos
- CNPJ Banestes mas expansível para todo SFN
- 01/2003 - 12/2022
- 37.920 observações (114.480 com nós de agregação)



# DADOS E VARIÁVEIS

## 4 METODOLOGIA

- Estrutura hierárquica
  - total
  - 4 mesorregiões
  - 13 microrregiões
  - 79 agências
- Estrutura agrupada
  - 2 modalidades de crédito



# PREVISÕES BASE

## 4 METODOLOGIA

- Previsões para fora da amostra a serem reconciliadas
- ETS (podem ser obtidas por qualquer método, incluindo ML)
- Treino: 01/2003 - 12/2021
- Horizonte de previsão ( $h$ ): 01/2022 - 12/2022, fora da amostra



# DESIGN DO EXPERIMENTO

## 4 METODOLOGIA

Dado um conjunto de previsões base, *o quanto mais precisas elas se tornam ao se aplicar um determinado método de reconciliação?*



# DESIGN DO EXPERIMENTO

## 4 METODOLOGIA

Metodologia de Spiliotis et al. (2021):

1. Previsões para dentro da amostra (ETS)
2. Treino dos modelos de ML para cada  $y_m$
3. Reconciliação ótima das previsões base (fora da amostra)
4. Agregação das previsões reconciliadas para obtenção dos demais níveis hierárquicos





# DESIGN DO EXPERIMENTO

## 4 METODOLOGIA

### METODOLOGIA ML VERSUS MINT

As principais diferenças entre os ambas metodologias são: (i) a utilização de algoritmos de ML ao invés de MQG (permitindo combinações não lineares), (ii) a não atribuição de peso de forma obrigatória para todos os nós da hierarquia e (iii) no ajuste de um modelo individual para cada série do nível mais desagregado, permitindo maior especialização e sendo capaz de se adaptar melhor aos diferentes padrões de cada série (SPILIOTIS et al., 2021).



# DESIGN DO EXPERIMENTO

## 4 METODOLOGIA

- Métodos de reconciliação:
  - Analíticos: *Bottom-up* e *MinT Shrink*
  - Regressão regularizada: *Elastic Net* (lasso, ridge e CV)
  - Árvores: *Random forest*, *XGBoost* e *LightGBM*
  - *Support Vector Machines* (SVM)
- Estratégias para obtenção das previsões *in-sample*:
  - *Rolling forecast*
  - *Reduced fitted base forecasts*
  - *Fitted base forecasts*



# ESTRATÉGIA ROLLING FORECAST

## 4 METODOLOGIA

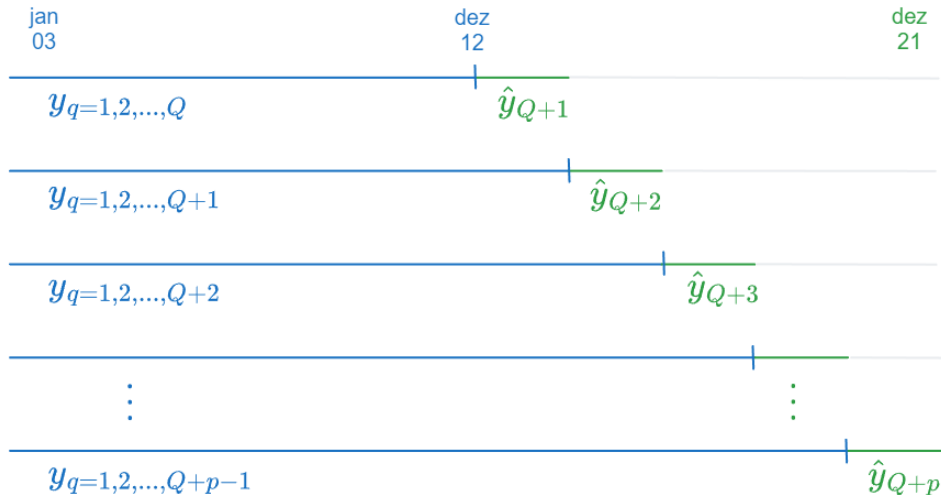


Figure 2: Estratégia *Rolling Forecast*



# ESTRATÉGIA REDUCED FITTED BASE

## 4 METODOLOGIA

- Desvantagem do *Rolling Forecast*: a cada  $h$ , um novo modelo é ajustado, totalizando  $m \times p = 17064$  modelos.
- *Reduced fitted base*: ajusta-se apenas  $m$  modelos, com dados até  $Q$  e então reestima-se esses modelos passando-se o restante do conjunto de treinamento ( $Q + p$ ), conservando os hiperparâmetros originais. Os valores ajustados dos modelos reestimados são utilizados para treinar os modelos de ML.



# ESTRATÉGIA FITTED BASE

## 4 METODOLOGIA

- Pode ser um problema obter previsões para dentro da amostra para séries de poucas observações ou incompletas
- *Fitted base*: utiliza-se os valores ajustados dos modelos de previsão base obtidos ao longo de todo o período  $q = 1, 2, 3, \dots, Q + p$  (todo conjunto de treinamento)



# OTIMIZAÇÃO DE HIPERPARÂMETROS

## 4 METODOLOGIA

- Bischl et al. (2021)
- $m$  modelos calibrados individualmente
- *grid* para *elastic net* e bayesiana os demais métodos
- *Resampling*: CV-10, RMSE



# SUMÁRIO

## 5 RESULTADOS

► CONTEXTUALIZAÇÃO DA PESQUISA

► OBJETIVOS

► RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS

► METODOLOGIA

► **RESULTADOS**

► REFERÊNCIAS



# ESTBAN - ANALÍTICOS

## 5 RESULTADOS

.model	agregado	mesorregiao	microrregiao	municipio	agencia	verbete	bottom	hierarquia
<b>RMSSE</b>								
base	0.197	0.690	0.846	<b>0.671</b>	0.717	0.183	0.656	0.657
bu	0.196	0.561	<b>0.706</b>	0.710	0.739	0.185	0.656	0.663
mint	0.172	0.619	0.722	0.689	<b>0.712</b>	<b>0.140</b>	<b>0.634</b>	<b>0.641</b>
<b>MASE</b>								
base	0.278	0.818	0.998	<b>0.790</b>	<b>0.886</b>	0.250	0.883	0.847
bu	0.240	0.572	<b>0.771</b>	0.820	0.895	0.221	<b>0.883</b>	0.844
mint	0.224	0.692	0.865	0.830	0.891	<b>0.164</b>	0.864	<b>0.837</b>





# ESTBAN - ML - ROLLING FORECAST

## 5 RESULTADOS

modelo	agregado	mesorregiao	microrregiao	municipio	agencia	verbete	bottom	hierarquia
<b>RMSSE</b>								
elastic net	0.280	0.763	1.178	1.211	1.251	0.770	1.062	1.094
lasso	0.196	0.726	1.054	0.995	1.043	0.501	0.839	0.882
lightgbm	1.407	1.628	1.530	1.294	1.322	0.883	0.972	1.095
random forest	1.227	1.397	1.303	1.118	1.171	0.725	0.858	0.949
ridge	0.416	0.776	1.131	1.511	1.535	0.919	1.357	1.368
svm	0.262	0.745	0.858	0.853	0.911	0.234	0.847	0.838
xgb	1.186	1.405	1.296	1.096	1.139	0.700	0.830	0.924
<b>MASE</b>								
elastic net	0.234	0.726	1.406	1.491	1.582	0.949	1.439	1.428
lasso	<u>0.166</u>	0.714	1.250	1.193	1.298	0.641	1.142	1.147
lightgbm	1.890	1.896	1.874	1.600	1.654	1.234	1.390	1.478
random forest	1.615	1.560	1.501	1.332	1.423	0.996	1.097	1.177
ridge	0.402	0.757	1.315	1.800	1.881	1.090	1.790	1.738
svm	0.306	0.684	0.862	0.991	1.108	0.290	1.251	1.143
xgb	1.542	1.564	1.479	1.297	1.373	0.948	1.022	1.115



# ESTBAN - ML - FITTED BASE

## 5 RESULTADOS

modelo	agregado	mesorregiao	microrregiao	municipio	agencia	verbete	bottom	hierarquia
<b>RMSSE</b>								
elastic net	0.777	0.986	1.086	1.038	1.149	0.579	0.924	0.961
lasso	0.661	0.955	1.074	0.909	1.008	0.530	0.826	0.862
lightgbm	1.495	1.649	1.557	1.300	1.342	0.923	0.999	1.116
random forest	1.204	1.397	1.294	1.098	1.150	0.694	0.839	0.930
ridge	1.001	1.146	1.247	1.208	1.327	0.689	1.125	1.147
svm	0.395	0.929	0.928	0.934	0.961	0.319	0.905	0.898
xgb	1.196	1.373	1.282	1.084	1.133	0.699	0.824	0.916
<b>MASE</b>								
elastic net	1.049	1.143	1.317	1.268	1.451	0.795	1.290	1.284
lasso	0.894	1.087	1.319	1.114	1.276	0.728	1.167	1.162
lightgbm	2.027	1.931	1.906	1.610	1.683	1.302	1.433	1.512
random forest	1.576	1.557	1.488	1.311	1.397	0.947	1.043	1.132
ridge	1.338	1.350	1.549	1.501	1.696	0.935	1.584	1.551
svm	0.445	0.947	1.080	1.163	1.226	0.341	1.282	1.217
xgb	1.545	1.509	1.476	1.293	1.375	0.942	1.017	1.109



# ESTBAN - ML - REDUCED FITTED BASE

## 5 RESULTADOS

modelo	agregado	mesorregiao	microrregiao	municipio	agencia	verbete	bottom	hierarquia
<b>RMSSE</b>								
elastic net	<b>0.090</b>	0.582	0.730	0.819	0.838	0.164	0.730	0.736
lasso	<u>0.132</u>	<b>0.523</b>	0.766	0.757	0.774	0.187	0.681	0.693
lightgbm	1.406	1.588	1.520	1.281	1.323	0.889	0.971	1.091
random forest	1.248	1.409	1.319	1.119	1.167	0.692	0.857	0.949
ridge	<u>0.141</u>	0.635	0.784	0.902	0.922	0.207	0.841	0.829
svm	0.187	0.743	0.767	0.792	0.834	0.295	0.807	0.792
xgb	1.218	1.347	1.253	1.084	1.140	0.708	0.844	0.927
<b>MASE</b>								
elastic net	<b>0.086</b>	0.584	0.834	0.973	1.008	0.208	0.944	0.922
lasso	<u>0.138</u>	<b>0.520</b>	0.883	0.907	0.933	0.216	0.891	0.878
lightgbm	1.879	1.831	1.832	1.580	1.652	1.236	1.388	1.470
random forest	1.636	1.576	1.526	1.333	1.418	0.947	1.065	1.155
ridge	<u>0.159</u>	0.630	0.894	1.047	1.087	0.231	1.073	1.021
svm	0.225	0.764	0.850	0.949	1.022	0.395	1.176	1.083
xgb	1.593	1.491	1.427	1.293	1.380	0.965	1.064	1.137



# ESTBAN - ML - TEMPO DE PROCESSAMENTO

## 5 RESULTADOS

Table 1: Resultados Estban: Tempo de processamento dos métodos de ML (em horas)

	xgb	random forest	elastic net	lasso	ridge	svm	lightgbm
reduced fitted base	19.235	5.483	1.428	0.879	0.993	1.279	3.290
fitted base	21.758	5.521	1.363	0.829	0.924	1.273	3.341
rolling forecast	20.908	5.429	1.345	0.838	0.929	1.285	3.377



# TOURISM - ANALÍTICOS

## 5 RESULTADOS

.model	agregado	state	region	purpose	bottom	hierarquia
<b>RMSSE</b>						
base	1.446	1.260	1.068	1.265	0.925	0.976
bu	2.580	1.634	1.113	2.004	0.925	1.011
mint	1.813	1.296	0.978	1.420	<b>0.876</b>	<b>0.923</b>
<b>MASE</b>						
base	1.533	1.399	1.132	1.330	0.979	1.036
bu	3.164	1.877	1.176	2.323	0.979	1.078
mint	2.086	1.449	1.021	1.512	<b>0.937</b>	<b>0.984</b>



# TOURISM - ML - ROLLING FORECAST

## 5 RESULTADOS

modelo	agregado	State	Region	Purpose	bottom	hierarquia
<b>RMSSE</b>						
elastic net	1.990	1.386	1.086	1.541	0.988	1.041
lasso	1.929	1.373	1.100	1.523	1.026	1.069
lightgbm	4.330	2.762	1.651	3.456	1.141	1.354
random forest	2.135	1.365	1.033	1.709	0.908	0.966
ridge	<u>1.256</u>	<u>1.185</u>	1.013	<u>1.202</u>	0.919	0.959
svm	<b>0.940</b>	<b>1.010</b>	1.076	<b>1.011</b>	1.100	1.097
xgb	2.340	1.451	1.114	1.892	0.964	1.031
<b>MASE</b>						
elastic net	2.360	1.572	1.145	1.653	1.058	1.115
lasso	2.264	1.557	1.168	1.593	1.110	1.155
lightgbm	5.505	3.214	1.763	4.060	1.200	1.448
random forest	2.579	1.528	1.073	1.816	0.961	1.020
ridge	<u>1.343</u>	<u>1.309</u>	1.058	<u>1.192</u>	0.981	1.020
svm	<b>1.070</b>	<b>1.096</b>	1.140	<b>1.033</b>	1.178	1.174
xgb	2.888	1.650	1.162	2.118	1.013	1.087



# TOURISM - ML - FITTED BASE

## 5 RESULTADOS

modelo	agregado	State	Region	Purpose	bottom	hierarquia
<b>RMSSE</b>						
elastic net	2.17	1.40	1.10	1.77	0.97	1.03
lasso	1.90	1.45	1.09	1.61	0.97	1.03
lightgbm	4.33	2.76	1.65	3.46	1.14	1.35
random forest	2.12	1.36	1.03	1.72	0.91	0.96
ridge	1.57	<u>1.16</u>	<b>0.97</b>	1.29	0.90	0.93
svm	1.50	<u>1.19</u>	1.05	1.38	1.04	1.05
xgb	2.27	1.42	1.10	1.83	0.96	1.02
<b>MASE</b>						
elastic net	2.59	1.57	1.16	1.95	1.04	1.10
lasso	2.21	1.67	1.16	1.73	1.04	1.11
lightgbm	5.50	3.21	1.76	4.06	1.20	1.45
random forest	2.54	1.54	1.07	1.84	0.97	1.02
ridge	1.77	<u>1.29</u>	<b>1.01</b>	<u>1.31</u>	0.96	0.99
svm	1.75	<u>1.30</u>	1.09	1.36	1.10	1.11
xgb	2.79	1.62	1.15	2.04	1.01	1.08



# TOURISM - ML - REDUCED FITTED BASE

## 5 RESULTADOS

modelo	agregado	State	Region	Purpose	bottom	hierarquia
<b>RMSSE</b>						
lightgbm	4.33	2.76	1.65	3.46	1.14	1.35
random forest	2.58	1.55	1.11	2.01	0.92	1.01
xgb	3.14	1.93	1.28	2.44	1.02	1.14
<b>MASE</b>						
lightgbm	5.50	3.21	1.76	4.06	1.20	1.45
random forest	3.18	1.76	1.17	2.24	0.98	1.07
xgb	3.93	2.27	1.38	2.86	1.07	1.21





# TOURISM - ML - TEMPO DE PROCESSAMENTO

## 5 RESULTADOS

	xgb	random forest	elastic net	lasso	ridge	svm	lightgbm
fitted base	15.767	3.687	1.540	1.282	1.371	1.977	3.535
rolling forecast	12.087	2.987	0.957	0.796	1.073	2.035	3.596
reduced fitted base	24.627	15.386	6.006	4.968	5.323	10.474	19.362



# SUMÁRIO

## 6 REFERÊNCIAS

- ▶ CONTEXTUALIZAÇÃO DA PESQUISA
- ▶ OBJETIVOS
- ▶ RECONCILIAÇÃO DE SÉRIES TEMPORAIS HIERÁRQUICAS
- ▶ METODOLOGIA
- ▶ RESULTADOS
- ▶ REFERÊNCIAS



# **BIBLIOGRAFIA**

## **6 REFERÊNCIAS**



# REFERÊNCIAS

## 6 REFERÊNCIAS



BECCALLI, E. et al. Earnings management, forecast guidance and the banking crisis. **The European Journal of Finance**, v. 21, n. 3, p. 242–268, fev. 2015. ISSN 1351-847X. DOI: 10.1080/1351847X.2013.809548. Disponível em: <<https://doi.org/10.1080/1351847X.2013.809548>>. Acesso em: 7 mai. 2023. Citado na p. 3.



BISCHL, B. et al. **Hyperparameter Optimization: Foundations, Algorithms, Best Practices and Open Challenges**. [S.l.]: arXiv, nov. 2021. arXiv:2107.05847 [cs, stat]. DOI: 10.48550/arXiv.2107.05847. Disponível em: <<http://arxiv.org/abs/2107.05847>>. Acesso em: 4 set. 2023. Citado na p. 38.



GROSS, C. W.; SOHL, J. E. Disaggregation methods to expedite product line forecasting. en. **Journal of Forecasting**, v. 9, n. 3, p. 233–254, 1990. ISSN 1099-131X. DOI: 10.1002/for.3980090304. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/for.3980090304>>. Acesso em: 18 jun. 2023. Citado na p. 18.



# REFERÊNCIAS

## 6 REFERÊNCIAS



HYNDMAN, R.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 3. ed. Melbourne, Austrália: OTexts, 2021. Disponível em: <<https://otexts.com/fpp3/>>. Citado na p. 5.



SEZER, O. B.; GUDELEK, M. U.; OZBAYOGLU, A. M. **Financial Time Series Forecasting with Deep Learning : A Systematic Literature Review: 2005-2019**. Turquia: arXiv, nov. 2019. arXiv:1911.13288 [cs, q-fin, stat]. Disponível em: <<http://arxiv.org/abs/1911.13288>>. Acesso em: 7 mar. 2023. Citado na p. 3.



SPILIOTIS, E. et al. Hierarchical forecast reconciliation with machine learning. en. **Applied Soft Computing**, v. 112, p. 107756, nov. 2021. ISSN 1568-4946. DOI: 10.1016/j.asoc.2021.107756. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494621006773>>. Acesso em: 11 jan. 2023. Citado nas pp. 15, 26, 32, 33.



# REFERÊNCIAS

## 6 REFERÊNCIAS



WICKRAMASURIYA, S. L.; ATHANASOPOULOS, G.; HYNDMAN, R. J. Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization. en. **Journal of the American Statistical Association**, v. 114, n. 526, p. 804–819, abr. 2019. ISSN 0162-1459, 1537-274X. DOI: 10.1080 / 01621459.2018.1448825. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/01621459.2018.1448825>>. Acesso em: 18 jun. 2023. Citado na p. 25.



# **MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA DE SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS DE INSTITUIÇÕES FINANCEIRAS**

DEFESA DE DISSERTAÇÃO DE MESTRADO

Alberson Miranda, PPGeco/UFES

Fevereiro de 2024