



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

2024



SUMÁRIO

1 CONTEXTUALIZAÇÃO DA PESQUISA

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ ABORDAGENS DE NÍVEL ÚNICO

▶ MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

▶ MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

▶ METODOLOGIA

▶ 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

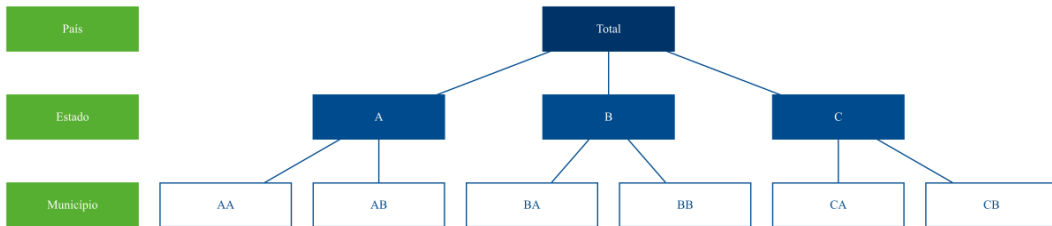
1 CONTEXTUALIZAÇÃO DA PESQUISA

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

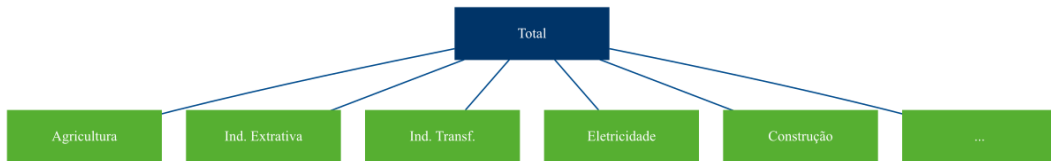
1 CONTEXTUALIZAÇÃO DA PESQUISA

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



SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

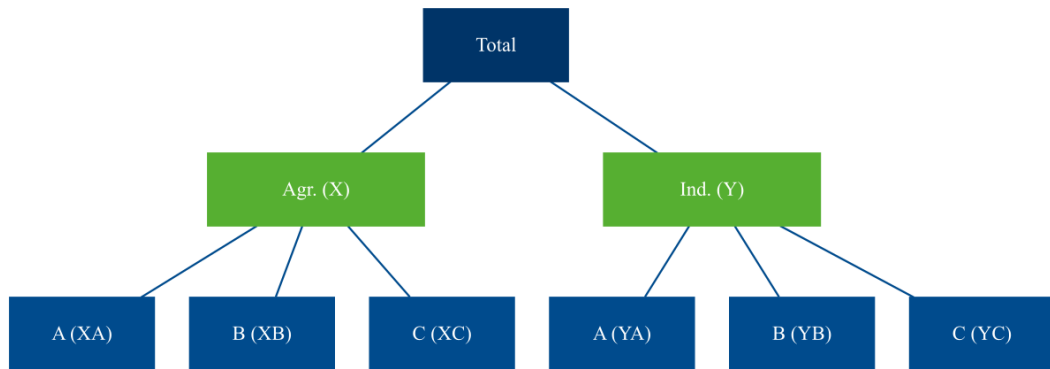
1 CONTEXTUALIZAÇÃO DA PESQUISA





SÉRIES TEMPORAIS HIERÁRQUICAS E AGRUPADAS

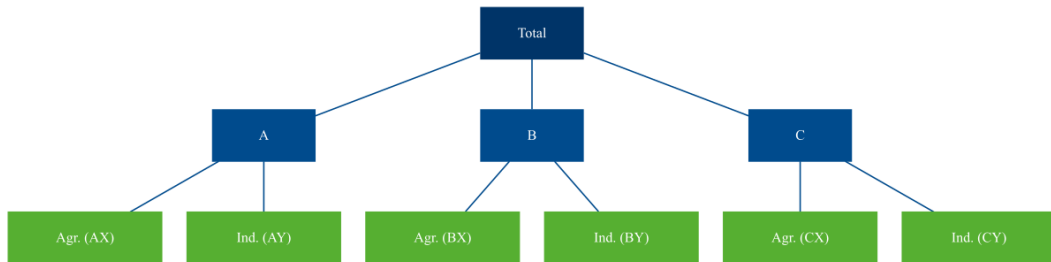
1 CONTEXTUALIZAÇÃO DA PESQUISA





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

► ABORDAGENS DE NÍVEL ÚNICO

► MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

► MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

► METODOLOGIA

► 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 variações metodológicas para a metodologia;
4. Obter ganhos de performance no contexto de séries temporais financeiras.



SUMÁRIO

3 ABORDAGENS DE NÍVEL ÚNICO

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ **ABORDAGENS DE NÍVEL ÚNICO**

▶ MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

▶ MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

▶ METODOLOGIA

▶ REFERÊNCIAS



BOTTOM-UP

3 ABORDAGENS DE NÍVEL ÚNICO

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



TOP-DOWN

3 ABORDAGENS DE NÍVEL ÚNICO

- 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 agrupadas¹.

¹Disambiguar 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 ABORDAGENS DE NÍVEL ÚNICO

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



SUMÁRIO

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ ABORDAGENS DE NÍVEL ÚNICO

▶ MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

▶ MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

▶ METODOLOGIA

▶ REFERÊNCIAS



NOTAÇÃO MATRICIAL

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

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

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

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

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

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

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

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

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

exemplo 3: matriz de reconciliação bottom-up



MATRIZ DE RECONCILIAÇÃO

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

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



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

PROBLEMA DE REGRESSÃO

- Minimização do traço da matriz de covariância dos erros das previsões reconciliadas (MinT^a)

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

- Mínimos Quadrados Generalizados (MQG)

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

^a(WICKRAMASURIYA; ATHANASOPOULOS; HYNDMAN, 2019).



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

- 1ª abordagem: $W_h = k_h I$
 - Mínimos Quadrados Ordinários (MQO)
 - Não há qualquer razão ou justificativa para supor homoscedasticidade
 - Erros descorrelacionados e equivariantes não fazem sentido em séries hierárquicas



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

- 2ª abordagem: $W_h = k_h \text{diag}(\hat{W}_1) \mid \hat{W}_1 = \frac{1}{T} \sum_{t=1}^T \hat{e}_{T+1} \hat{e}_{T+1}'$
 - Mínimos Quadrados Ponderados (MQP)
 - $\hat{e}_{T+1} = y_{T+1} - \hat{y}_{T+1}$
 - Erros descorrelacionados e ponderados pela variância do erro das previsões base



RECONCILIAÇÃO ÓTIMA

4 MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

- 3ª abordagem: $W_h = k_h \hat{W}_1$
 - MQG irrestrito
 - Quando $m \geq T$, pode-se usar estimador de encolhimento
 - Leva em consideração correlação e variância ao longo de toda estrutura



SUMÁRIO

5 MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

▶ CONTEXTUALIZAÇÃO DA PESQUISA

▶ OBJETIVOS

▶ ABORDAGENS DE NÍVEL ÚNICO

▶ MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA

▶ MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

▶ METODOLOGIA

▶ REFERÊNCIAS



ELASTIC NET

5 MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

- O *elastic net* (ZOU; HASTIE, 2005) é um método de regressão regularizada que combina as penalidades *lasso* e do *ridge*
- A função objetivo a ser minimizada é dada por

$$L(\lambda_1, \lambda_2, \beta) = \|\mathbf{y} - \mathbf{X}\beta\|^2 + \lambda_2 \|\beta\|^2 + \lambda_1 \|\beta\|_1 \quad (7)$$



ELASTIC NET

5 MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA

- Solução dada por

$$\hat{\beta} = \arg \min_{\beta} |\mathbf{y} - \mathbf{X}\beta|^2, \text{ sujeito a } (1 - \alpha)|\beta|_1 + \alpha|\beta|^2 \leq t \quad (8)$$

- $\alpha = \frac{\lambda_2}{\lambda_1 + \lambda_2}$ e $t \in \mathbb{R}^+$ ($t = L_1$)
- $\alpha = 0 \implies \lambda_2 = 0$, *lasso*
- $\alpha = 1 \implies \lambda_1 = 0$, *ridge*



SUMÁRIO

6 METODOLOGIA

- ▶ CONTEXTUALIZAÇÃO DA PESQUISA
- ▶ OBJETIVOS
- ▶ ABORDAGENS DE NÍVEL ÚNICO
- ▶ MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA
- ▶ MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA
- ▶ **METODOLOGIA**
- ▶ REFERÊNCIAS



DADOS E VARIÁVEIS

6 METODOLOGIA

- Estatística Bancária Mensal e por Município (Bacen)
- Municípios (IBGE)
- Saldos de crédito
- CNPJ Banestes (expansível para todo SFN)
- 01/2010 - 02/2023
- 43.923 observações



DADOS E VARIÁVEIS

6 METODOLOGIA

- Estrutura hierárquica
 - total
 - 4 mesorregiões
 - 13 microrregiões
 - 96 agências
- Estrutura agrupada
 - 3 modalidades de crédito



DESIGN DO EXPERIMENTO

6 METODOLOGIA

- \hat{y} via Arima ou ETS
- *benchmark* (RMSE, tempo de processamento):
 - *Top-down*
 - *bottom-up*
 - MinT OLS
 - MinT WLS
 - MinT *Shrink* ($t = 156, m = 96 \times 3$)
 - *Elastic Net* Lasso
 - *Elastic Net* Ridge
 - *Elastic Net* CV-5



DESIGN DO EXPERIMENTO

6 METODOLOGIA

- Procedimentos para o método *elastic net* (SPILIOTIS et al., 2021)
 1. Obter \hat{y}_{t+1}
 2. Treinar modelo *elastic net* usando (1) como preditores e y_{t+1} como alvo
 3. Obter \hat{y}_{T+h}
 4. Obter \tilde{y}_{T+h} passando os preditores (3) ao modelo treinado em (2)
 5. Agregar as previsões reconciliadas para obter as previsões nos demais níveis hierárquicos



SUMÁRIO

7 REFERÊNCIAS

- ▶ CONTEXTUALIZAÇÃO DA PESQUISA
- ▶ OBJETIVOS
- ▶ ABORDAGENS DE NÍVEL ÚNICO
- ▶ MÉTODOS ANALÍTICOS PARA RECONCILIAÇÃO ÓTIMA
- ▶ MÉTODOS DE MACHINE LEARNING PARA RECONCILIAÇÃO ÓTIMA
- ▶ METODOLOGIA
- ▶ REFERÊNCIAS



BIBLIOGRAFIA

7 REFERÊNCIAS



REFERÊNCIAS

7 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.



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.



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



REFERÊNCIAS

7 REFERÊNCIAS



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, 37.



REFERÊNCIAS

7 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. 26.



ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. en. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 67, n. 2, p. 301–320, 2005. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2005.00503.x>. ISSN 1467-9868. DOI: 10.1111 / j.1467 - 9868.2005.00503.x. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2005.00503.x>>. Acesso em: 5 jun. 2023. Citado na p. 31.



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

2024