

Exploração Visual e Computacional de Dados Clínicos para Monitoramento de Indicadores Cardiovasculares

Guilherme P. Menezes¹, Miriam A. Z. Marcolino², Luciana R. de Lara²,
Carisi A. Polanczyk², Carla M.D.S Freitas¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Av. Bento Gonçalves, 9500 – 91.501-970 – Porto Alegre – RS – Brazil

²IATS - Instituto de Avaliação de Tecnologia em Saúde (UFRGS-HCPA)
Rua Ramiro Barcelos, 1981 – 90035-006 – Porto Alegre – RS

{guilherme.menezes, carla}@inf.ufrgs.br

{mmarcolino, lrlara, cpolanczyk}@hcpa.edu.br

Abstract. *This study integrates visual analytics and machine learning to monitor indicators in cardiovascular care. Initially, we are investigating two indicators: in-hospital mortality and 30-day readmission. Using data from hospital admissions, cardiovascular procedures, and ICD-10 diagnoses, we performed feature engineering and trained Random Forest and XGBoost models. We observed substantial gains for readmission after oversampling (recall up to 0.98), whereas mortality remains more difficult to predict. We present preliminary results and discuss limitations, biases, and next steps.*

Resumo. *Este estudo integra exploração visual e aprendizado de máquina para monitorar indicadores na assistência cardiovascular. Inicialmente, estamos investigando dois indicadores: mortalidade hospitalar e reinternação em 30 dias. A partir de dados de internações, procedimentos cardiovasculares e CID-10, realizamos engenharia de atributos e treinamos Random Forest e XGBoost. Houve ganho expressivo para reinternação após oversampling (recall até 0,98), ao passo que a mortalidade apresenta maior dificuldade preditiva. Apresentamos resultados preliminares e discutimos limitações, vieses e próximos passos.*

1. Introdução

Doenças cardiovasculares seguem entre as principais causas de internação e óbito, pressionando gestores a acompanhar, com agilidade e rigor, diversos indicadores [Bertoletti 2023, Oliveira and et al. 2024]. Embora hospitais acumulem grandes volumes de dados clínico-administrativos, transformá-los em informação acionável permanece desafiador devido a bases fragmentadas, heterogeneidade de codificações e ausência de *pipelines* reprodutíveis que conectem análise exploratória, predição e uso gerencial [Silva and Sanine 2020, de Lara et al. 2024].

Nas últimas décadas, modelos de aprendizado de máquina vêm sendo aplicados a tarefas afins (p.ex., previsão de eventos adversos e altas), mas persistem limitações: (i) dependência de variáveis pouco disponíveis na rotina; (ii) desbalanceamento severo das classes, que distorce métricas; (iii) generalização frágil quando se misturam perfis clínicos distintos; e (iv) baixa integração com exploração visual para inspeção de cortes e comunicação de resultados a tomadores de decisão [Talukdar et al. 2023, Thirunavukarasu et al. 2022, Wanderer et al. 2016].

A carência de soluções automatizadas que facilitem o acesso e a compreensão desses dados reforça a necessidade de estudos que, em um mesmo fluxo reprodutível, unam (a) análise visual orientada a indicadores hospitalares e (b) modelagem preditiva focada especificamente em mortalidade hospitalar e reinternação em 30 dias na assistência cardiovascular, explorando de forma explícita os procedimentos cardiovasculares ao lado de variáveis demográficas/administrativas e mapeamentos CID-10, com atenção sistemática ao desbalanceamento e à interpretabilidade prática dos resultados.

Este artigo apresenta um estudo em andamento que integra: (i) criação de um dataset clínico-administrativo; (ii) um dashboard interativo para exploração visual e monitoramento; e (iii) uma linha de base preditiva com Random Forest e XGBoost, avaliando estratégias de balanceamento (p.ex., *oversampling/undersampling*) e treinos estratificados por grupo de procedimento. O foco aqui é discutir o desenho experimental e as métricas apropriadas sob desbalanceamento, oferecendo evidências e diretrizes para uso gerencial.

As principais contribuições são: (1) um *pipeline* reprodutível que conecta exploração visual, preparação de dados e predição dos dois indicadores (reinternação em 30 dias e mortalidade hospitalar); (2) a estratificação por procedimento, evidenciando heterogeneidade de risco e impacto na performance dos modelos; (3) o tratamento explícito do desbalanceamento, com avaliação por *precision-recall* e discussão de implicações práticas para triagem de casos; e (4) um protótipo de *dashboard* para comunicação com *stakeholders* e priorização de pacientes/coortes.

2. Materiais e Métodos

Utilizamos uma base clínico-administrativa de internações hospitalares em Cardiologia do Hospital de Clínicas de Porto Alegre (HCPA), contendo múltiplas internações por paciente. O conjunto foi obtido anonimizado e incluiu campos como: PRONTUARIO_ANONIMO, datas de internação/alta/cirurgia/óbito, idade, sexo, escolaridade, código SIGTAP (Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS), grupo de procedimento derivado, além das variáveis-alvo MORTE_HOSPITALAR e REINTERNADO_30DIAS.

Os critérios de inclusão adotados foram: internações com procedimento cardiovascular mapeado à SIGTAP, diagnóstico CID-10, datas válidas e desfechos observáveis. Foram excluídos registros com inconsistências temporais (alta antes da admissão), duplicidades pós-anonimização e transferências sem desfecho final.

A limpeza de dados envolveu padronização de datas e códigos, tratamento de ausências (mediana para numéricos; categoria “desconhecido” para nominais) e derivação de atributos (tempo de internação, tempo até cirurgia, dummies para CID-10 e grupo de procedimento, além de interações simples, como, por exemplo, idade \times procedimento). Variáveis redundantes foram eliminadas por VIF e correlação de Spearman.

O particionamento foi feito por grupo de procedimento, com treino/teste 70/30 (estratificação pelos desfechos) e validação cruzada 5-fold no treino. Avaliamos Random Forest (RF) [Breiman 2001] e XGBoost (XGB) [Chen and Guestrin 2016], ajustados via *random search*. As variáveis categóricas foram codificadas por *one-hot encoding* e as numéricas, quando apropriado, padronizadas por z-score ($z = (x - \mu)/\sigma$). Para o desbalanceamento, aplicamos **SMOTE** (*Synthetic Minority Over-sampling Technique*) na reinternação e testamos *undersampling* e *class weights* na mortalidade; a amostragem ocorreu apenas nos *folds* de treino, evitando vazamento.

As métricas relatadas incluíram acurácia, precisão, *recall*, F1, AUROC e PR-AUC, privilegiando *recall* e PR-AUC. Curvas ROC/PR, matrizes de confusão e intervalos *bootstrap* (1.000 reamostragens) foram usados para comparação. Também avaliamos modelagem estratificada por grupo de procedimento (quando havia volume suficiente de eventos) e construímos um *dashboard* prototipado em Figma¹ para visualização gerencial dos resultados, incluindo séries temporais, *rankings* de risco, filtros por coorte e explicabilidade dos modelos.

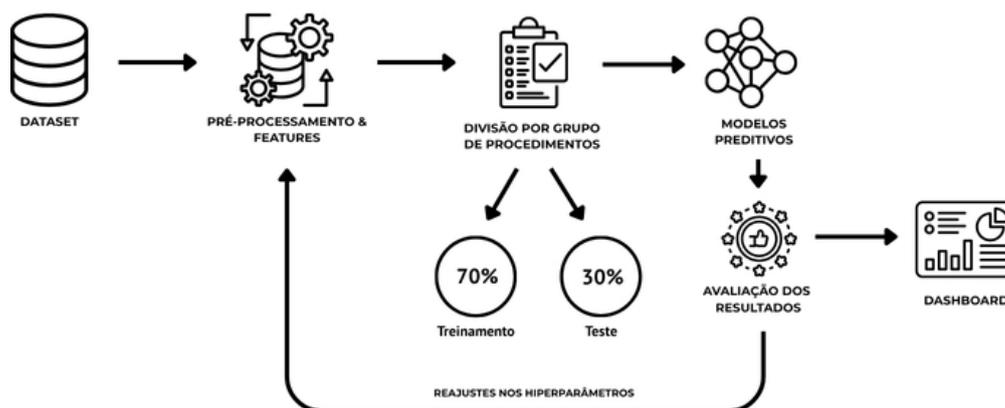


Figura 1. Pipeline desenvolvido

3. Resultados e Discussão

Em relação ao panorama da coorte, a base apresenta baixa prevalência de eventos (reinternação em 30 dias em poucos pontos percentuais; mortalidade hospitalar rara), reforçando a necessidade de métricas robustas ao desbalanceamento (p.ex., PR-AUC, *recall*).

No caso da variável “Reinternação em 30 dias”, os modelos tiveram *recall* muito baixo para a classe positiva, sem reamostragem. Com *oversampling* (SMOTE) aplicado no treino/teste, houve ganho substancial de sensibilidade. XGBoost atingiu *recall* $\approx 0,98$ (com leve queda de precisão) e Random Forest manteve *recall* alto ($\approx 0,95$) com precisão competitiva. O PR-AUC aumentou de forma consistente, enquanto a AUROC variou pouco, padrão típico em bases desbalanceadas. A implicação prática é que priorizar *recall* é desejável em triagem (não perder casos que merecem intervenção), mitigando falsos positivos por regras de priorização.

No que se refere à variável “Mortalidade hospitalar”, mesmo com *undersampling* e *class weights* (atribuição de pesos às classes na função de perda, aumentando a penalização para a classe positiva rara), o desempenho foi moderado (*recall*/ F_1 limitados), coerente com a prevalência menor, heterogeneidade clínica e ausência de sinais temporais mais finos. Leitura de erros: Falsos negativos concentraram-se em evoluções rápidas; falsos positivos, em perfis de alto risco basal com boa resposta terapêutica.

Analisando a estratificação por procedimento (GRUPO_PROCEDIMENTO), observa-se que, em grupos com amostra suficiente, a estratificação elevou o F1 para

¹<https://www.figma.com/>

reinternação (faixa 0,7+). Em grupos com baixa amostragem, manteve-se o modelo global para reduzir a variância. Para mortalidade, os ganhos foram heterogêneos, sugerindo coortes específicos (p.ex., revascularização vs. angioplastia) e mais dados.

4. Conclusões

O estudo mostrou que a integração de exploração visual e aprendizado de máquina é viável para monitorar indicadores cardiovasculares críticos. Os resultados indicam que a reinternação em 30 dias pode ser prevista com bom desempenho após técnicas de balanceamento, enquanto a mortalidade hospitalar permanece um desafio diante da baixa prevalência e heterogeneidade clínica. O *dashboard* prototipado reforça o potencial de aplicação prática, ao traduzir os achados em ferramentas de análise visual. Como próximos passos, destacam-se a ampliação da base de dados, com a incorporação de atributos temporais e textuais, teste e validação de outros algoritmos e desenvolvimento pleno do *dashboard*, visando consolidar uma solução confiável para o uso hospitalar.

Agradecimentos

O presente trabalho é apoiado pela CAPES (Finance Code 001), pelo CNPq (IATS), pela FAPERGS (RITE CIARS, proc. 22/2551-0000390-7) e pelo CI-IA Saúde (proc. 23072.229619/2023-51) e aprovado pelo Comitê de Ética do HCPA (AGHUSe Pesquisa 2023-0086, CAAE 68785923.0.1001.5327).

Referências

- Bertoletti, O. A. (2023). Boas práticas em cardiologia – uma lição a partir dos indicadores de desempenho. *Arquivos Brasileiros de Cardiologia*, 120(2):e20230033.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA. ACM.
- de Lara, L. R., Marcolino, M. A. Z., Schneider, N. B., Casiraghi, Y., Godoy, L., Aguiar, R. A. D. S., De Bairos, J. E. D. A., Etges, A. P. B. D. S., and Polanczyk, C. A. (2024). Performance indicators of cardiovascular care: a scoping review protocol. *OSF Registries*.
- Oliveira, G. M. M. and et al. (2024). Estatística Cardiovascular – Brasil 2023. *Arquivos Brasileiros de Cardiologia*, 121(2):e20240079.
- Silva, M. A. and Sanine, P. R. (2020). Interoperabilidade entre os sistemas de informação em saúde brasileiros: uma revisão integrativa. *Revista de Saúde Pública de Mato Grosso do Sul*, 3(2):17–29.
- Talukdar, J., Singh, T. P., and Barman, B. (2023). *Artificial Intelligence in Healthcare Industry*. Springer.
- Thirunavukarasu, R., C, G. P. D., R, G., Gopikrishnan, M., and Palanisamy, V. (2022). Towards computational solutions for precision medicine based big data healthcare system using deep learning models: A review. *Computers in Biology and Medicine*, 149:106020.
- Wanderer, J. P., Nelson, S. E., Ehrenfeld, J. M., and Sandberg, W. S. (2016). Clinical data visualization: The current state and future needs. *Journal of Medical Systems*, 40(11):275.