

FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

UNIVERSIDADE DE SÃO PAULO

Uso de algoritmos de aprendizado de máquina para o diagnóstico precoce de infecção da corrente sanguínea em pacientes internados na unidade de terapia intensiva pediátrica

Aluno: Felipe Martins Liporaci

Orientadora: Profa. Dra. Ana Paula de Carvalho Panzeri Carlotti

Pesquisador colaborador: Danilo Panzeri Carlotti

Ribeirão Preto

2021

INTRODUÇÃO

A infecção hospitalar é responsável por uma significativa parcela da morbimortalidade de pacientes em Centro de Terapia Intensiva- Pediátrico (CTIP). Dentre os sítios de infecção hospitalar, a infecção de corrente sanguínea (ICS) é responsável por cerca de 30% da mortalidade total em contexto de CTIP [1, 2]. A detecção precoce de ICS é imprescindível para que haja intervenção adequada rapidamente, evitando desfechos desfavoráveis [3, 4].

Entretanto, a habilidade dos médicos em preverem ICS pode ser limitada, o que, conseqüentemente, pode atrasar o diagnóstico e o tratamento adequado. Existem modelos que propõem antever a ocorrência de ICS, porém não têm aplicação clínica, visto que nenhum desses modelos é específico para pacientes internados em CTIP [5].

Como a detecção precoce é imprescindível e não existem modelos para prever a existência de ICS em pacientes internados em CTIP, é importante investigar métodos que possibilitem o uso de ferramentas para diagnosticar precocemente o problema, a fim de instituir o tratamento em momento oportuno, visando à melhora da morbimortalidade.

Aprendizado de máquina ou *machine learning* (ML) é uma ferramenta matemática e estatística que permite que, por meio de uma grande quantidade de dados, a máquina aprenda a resolver tarefas complexas sem que tenha sido explicitamente programada para tal. Apesar de existirem diversas publicações usando modelos de ML na área da medicina, poucos projetos tiveram aplicação prática importante.

OBJETIVOS

O presente estudo tem por objetivo desenvolver um algoritmo baseado em aprendizado de máquina ou *machine learning* (ML) para identificar pacientes com alto risco de ICS, avaliando dados demográficos, clínicos e laboratoriais coletados rotineiramente.

HIPÓTESE DO ESTUDO

A hipótese do estudo é que a utilização de modelo de ML possibilite o diagnóstico precoce e o tratamento em tempo oportuno de ICS em pacientes internados na unidade de terapia intensiva pediátrica.

CASUÍSTICA E MÉTODOS

Desenho do estudo

Trata-se de estudo de coorte retrospectivo, com coleta de dados de prontuário eletrônico de pacientes internados no CTIP do HC Criança do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (HCFMRP-USP).

Aspectos éticos

O estudo será submetido à apreciação do Comitê de Ética em Pesquisa do HCFMRP-USP e será solicitada dispensa do termo de consentimento livre e esclarecido, pela natureza retrospectiva do estudo.

Critérios de inclusão

Serão elegíveis para o estudo todos os pacientes de 0-18 anos de idade internados no CTIP do HC Criança do HCFMRP-USP no período de 01 de janeiro a 31 de dezembro de 2019, que tiveram pelo menos um resultado de hemocultura positiva durante a internação.

Critérios de exclusão

Serão excluídos pacientes com resultado positivo de hemocultura, com crescimento de germe contaminante.

Fonte de dados

Os dados serão obtidos da base de dados digital do sistema de prontuários eletrônicos do HCFMRP-USP chamado Athos. Serão coletados dados demográficos, clínicos, laboratoriais e de desfecho (alta ou óbito).

Coleta de dados

Serão coletados os seguintes dados dos prontuários eletrônicos dos pacientes:

- Dados demográficos: idade, sexo, peso, estatura, índice de massa corporal.
- Comorbidades: desnutrição/ obesidade, diabetes, imunodeficiência congênita ou adquirida, infecção por HIV, leucemia, linfomopatologias oncológicas, pulmonares, gastroenterológicas ou cardíacas. , neutropenia
- Escore de gravidade PRISM III nas primeiras 4 horas após admissão no CTIP.
- Escore PELOD nas primeiras 24 e 48 horas após internação no CTIP.
- Uso de dispositivos invasivos (cateter venoso central, cateter venoso periférico, cateter arterial, sonda vesical de demora, cânula traqueal) e tempo de uso.
- Dados laboratoriais: valores de proteína C-reativa (PCR), hemograma com contagens globais e diferenciais de leucócitos.
- Dados clínicos: frequência cardíaca, frequência respiratória, pressão arterial, temperatura, saturação de oxigênio.
- Dados de tratamento: Nutrição parenteral total e tempo de uso, período pós-operatório, transplante de medula óssea ou órgãos sólidos, uso de quimioterapia.

Processamento de dados

Os dados coletados de todos os pacientes serão, inicialmente, pré-processados para servir a uma análise exploratória para entender se há padrões reconhecíveis pela equipe médica nos valores encontrados antes e posterior ao início da infecção, detectada pelos exames referência.

Há duas alternativas que serão testadas para o desenvolvimento de modelos de auxílio à tomada de decisão dos médicos.

A primeira delas é considerar as variações de exames e aferições como séries temporais. Isso é feito para usar algoritmos como VAR (*Vector Autoregression Model*) [6] para modelar como as variáveis se comportam e como se influenciam no tempo. Modelos de série temporal tendem, contudo, a exigir mais processamento computacional e uma quantidade de observações maior.

A segunda abordagem é construir modelos preditivos usando modelos não lineares [7, 8] como KNN (K nearest neighbors), Regressão logística, Gradiente boosting trees e comitês de máquinas, que ensinam a vários modelos diferentes a modelar o problema e posteriormente criam um “comitê” que vota pelo outcome, havendo então um outro modelo que faz o aprendizado das estimativas desses vários modelos para tomar sua decisão. Serão testadas duas estratégias. A primeira estratégia é usar estes modelos para classificar os pacientes em pacientes com e sem infecção. A segunda estratégia será ensinar os modelos o tipo de variação “atípica” que indica uma infecção, sendo o baseline apreendido do paciente o único referencial para análise

Proposta de Cronograma

ETAPAS DA PESQUISA	PERÍODO DA PESQUISA
Período de análise pelo CEP	Meses 01 e 02
Permissão da utilização dos dados	Meses 03 e 04
Coleta de dados	Meses 05 a 08
Análise de dados	Meses 09 a 12
Redação do manuscrito para publicação dos dados	Meses 13 a 16
Elaboração e submissão do relatório final	Meses 17 a 24

REFERÊNCIAS

1. Wisplinghoff H, Bischoff T, Tallent SM et al (2004) Nosocomial blood- stream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. Clin Infect Dis 39:309–317. <https://doi.org/10.1086/421946>
2. Savage RD, Fowler RA, Rishu AH et al (2016) The effect of inadequate initial empiric antimicrobial treatment on mortality in critically ill patients with bloodstream

infections: a multi-centre retrospective cohort study. PLoS One 11:e0154944.

<https://doi.org/10.1371/journal.pone.0154944>

3. Weiss SL, Peters MJ, Alhazzani W et al (2020) Executive Summary: Surviving Sepsis Campaign International Guidelines for the Management of Septic Shock and Sepsis-Associated Organ Dysfunction in Children. *Pediatr Crit Care Med*. 2020;21(2):186
4. Weiss SL, Peters MJ, Alhazzani W et al (2020) Surviving Sepsis Campaign International Guidelines for the Management of Septic Shock and Sepsis-Associated Organ Dysfunction in Children. *Pediatr Crit Care Med*. 2020;21(2):e52
5. Eliakim-Raz N, Bates DW, Leibovici L (2015) Predicting bacteraemia in validated models-a systematic review. *Clin Microbiol Infect* 21:295–301.
<https://doi.org/10.1016/j.cmi.2015.01.023>
6. Bose, Eliezer¹; Hravnak, Marilyn²; Clermont, Gilles³ 283, *Critical Care Medicine*: December 2014 - Volume 42 - Issue 12 - p A1428-A1429 doi: 10.1097/01.ccm.0000457780.29828.f7
7. Alanazi, H.O., Abdullah, A.H., Qureshi, K.N. *et al*. Accurate and dynamic predictive model for better prediction in medicine and healthcare. *Ir J Med Sci* **187**, 501–513 (2018). <https://doi.org/10.1007/s11845-017-1655-3>
8. Steven C Bagley, Halbert White, Beatrice A Golomb, Logistic regression in the medical literature:: Standards for use and reporting, with particular attention to one medical domain, *Journal of Clinical Epidemiology*, Volume 54, Issue 10, 2001, Pages 979-985, ISSN 0895-4356, [https://doi.org/10.1016/S0895-4356\(01\)00372-9](https://doi.org/10.1016/S0895-4356(01)00372-9).