

O que todo intensivista deveria saber sobre *Big Data* e aprendizado da máquina na unidade de terapia intensiva

What every intensivist should know about Big Data and targeted machine learning in the intensive care unit

A crescente importância dos *Big Data* na saúde

A conjugação do acesso cada vez maior a grandes quantidades de dados clínicos com o substancial progresso obtido em relação ao aprendizado da máquina (AM) e a inteligência artificial (IA) criou novas e imprevistas oportunidades para a ciência de dados na saúde. *Big Data* é descrito como possuindo pelo menos três características distintas: volume, velocidade e variedade. Porém, com relação à saúde, inclui também variabilidade e valor.⁽¹⁾ Assim, é um grande desafio extrair alguma informação útil dos *Big Data* com a utilização dos métodos estatísticos tradicionais.⁽²⁾ A análise dos *Big Data* tem um imenso potencial para melhoria da qualidade dos cuidados, ajudando os profissionais de saúde na obtenção de decisões mais personalizadas, redução de perdas e erros, e a possibilidade de diminuição do custo dos cuidados.⁽³⁾ A previsão da disfunção de um órgão antes que ela ocorra pode ser extremamente útil para (i) tomar decisões terapêuticas melhores e mais individualizadas e (ii) em alguns casos, prevenir a ocorrência da falência do órgão, por meio de ajuste apropriado e antecipado da terapêutica. Além disto, a capacidade de prever uma futura deterioração pode ser muito útil para ajudar a liderança clínica em uma alocação proativa de recursos humanos. Malak et al. recentemente propuseram uma arquitetura de gestão de risco por múltiplos agentes com base em *Big Data* e análise para criar um ambiente colaborativo e em tempo real para o cuidado de neonatos com condições críticas na unidade de terapia intensiva (UTI) neonatal.⁽⁴⁾

Fontes de *Big Data* em saúde

A “revolução dos dados” na saúde e, em última instância, na terapia de pacientes críticos, depende da capacidade de dar fluxo e armazenamento a grandes quantidades de informação em um repositório protegido encriptado e central. Prontuários médicos eletrônicos, monitores junto ao leito, dispositivos de administração de fármacos, ventiladores ou máquinas de diálise geram dados continuamente. Vem se tornando possível combinar estes dados com resultados de exames laboratoriais, procedimentos, notas de cuidadores, relatórios de imagem e, finalmente, com desfechos, inclusive aqueles em longo prazo referentes a aspectos funcionais e comportamentais. Por exemplo, a *Mayo Clinic* desenvolveu um armazenamento de dados denominado *Multidisciplinary Epidemiology and Translational Research in Intensive Care Data Mart* (METRIC),⁽⁵⁾ enquanto o *Beth Israel Deaconess Medical Center* (BIDMC) lançou uma grande base de dados similar, que compreende dados anônimos relacionados à saúde de mais de 40 mil pacientes que estiveram internados em UTIs, o MIMIC-III.⁽⁶⁾ Ambas as

1. Institut National de la Santé et de la Recherche Médicale - Paris, France.

2. Université de Paris - Paris, France.

3. The ACTERREA Research Group - Paris, France.

4. Department of Anesthesia and Perioperative Medicine, Zuckerberg San Francisco General Hospital, University of California San Francisco - California, United States.

5. Department of Anesthesia and Intensive Care Medicine, Hôpital Européen Georges Pompidou - Paris, France.

Conflitos de interesse: Nenhum.

Submetido em 1º de fevereiro de 2019

Aceito em 5 de maio de 2019

Autor correspondente:

Romain Pirracchio

Department of Anesthesia and Perioperative Medicine

Zuckerberg San Francisco General Hospital

University of California San Francisco

1001 Potrero avenue

CA94110 San Francisco

United States

E-mail: romain.pirracchio@ucsf.edu

DOI: 10.5935/0103-507X.20190069

bases de dados estão amplamente disponíveis para fins de pesquisa científica. Cada vez mais UTIs, centros médicos e mesmo redes de saúde em grande escala vêm desenvolvendo soluções para o armazenamento e a análise de dados referentes a pacientes e a comparação entre diferentes sistemas e organizações.⁽⁷⁾

Aprendizado da máquina para análise preditiva e suporte à decisão na unidade de terapia intensiva

Porque “os *Big Data* incluem observações heterogêneas, multiespectrais, incompletas e imprecisas derivadas de diferentes origens”,⁽⁸⁾ é necessário que se desenvolvam análises e inferências analíticas apropriadas. O AM, que é o componente da IA que permite aos computadores fazer escolhas e previsões dirigidas por dados, é hoje considerado a solução de escolha para domar os *Big Data* em medicina.⁽⁹⁾ Obviamente, o AM tem a capacidade de modelar relacionamentos complexos entre grandes aspectos explanatórios e os desejados resultados, como desfechos de pacientes. Os algoritmos de AM são, em geral, divididos em duas categorias diferentes: métodos paramétricos *versus* métodos não paramétricos, algoritmos supervisionados *versus* não supervisionados, e algoritmos individuais *versus* algoritmos montados (Figura 1). Algoritmos de aprendizado supervisionados são utilizados

para descobrir o relacionamento entre fatores explanatórios potenciais e um ou mais desfechos selecionados. São comumente aplicados em terapia intensiva para predição de eventos clínicos, como mortalidade na UTI.⁽¹⁰⁾ Em algoritmos de aprendizado não supervisionados não há um desfecho selecionado específico; o objetivo é, essencialmente, escavar profundamente a estrutura de dados para identificar a correlação entre fenômenos e criar grupos de características. Estes algoritmos são, atualmente, utilizados principalmente na medicina de precisão, na qual o objetivo é descobrir subgrupos de pacientes que compartilham as mesmas características clínicas ou moleculares.⁽¹¹⁾

Perspectivas

A *Food and Drug Administration* (FDA) descreve medicina de precisão como a que proporciona “ao paciente certo, o fármaco certo, na dose certa e no momento certo”.⁽¹²⁾ Com o desenvolvimento de novos algoritmos de AM, vem se tornando viáveis, em um futuro visível analisar em tempo real, quantidades gigantescas de dados diretamente, junto ao leito, para proporcionar previsões mais personalizadas e relevantes. Este campo da análise de fluxo, no qual os dados são colhidos e utilizados em sequência para atualizar os algoritmos preditos atuais, é referido

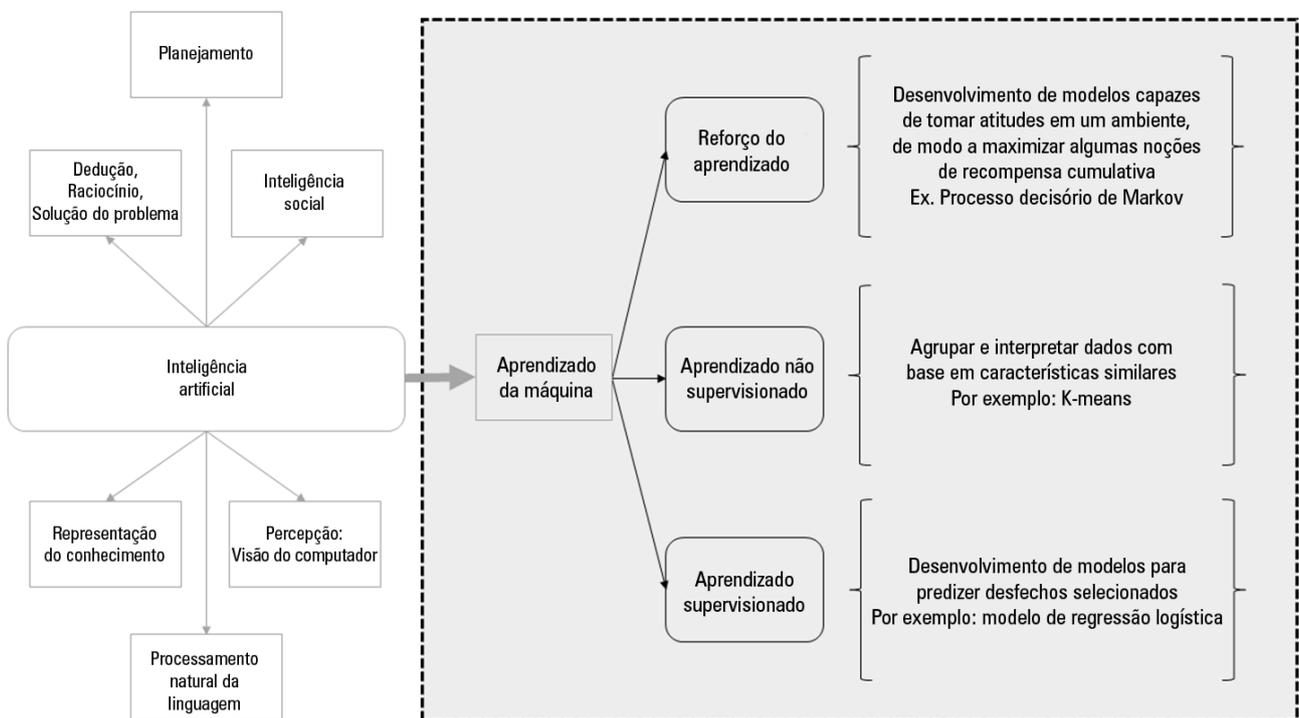


Figura 1 - Inteligência artificial e diferentes tipos de algoritmos de aprendizado da máquina.

como AM *online*.⁽⁷⁾ Uma tecnologia assim automatizada, utilizável junto ao leito, é o caminho para o alvo principal da medicina de precisão. Assim, os próximos desafios serão criar ferramentas de suporte em tempo real para tomada de decisão personalizada, o que permite ao médico adaptar melhor seu tratamento para pacientes em condições críticas. Esta abordagem atual, denominada análise prescritiva, refere-se à predição do efeito de tratamentos em nível do paciente. Uma abordagem estatística derivada de métodos de inferência causal pode ser utilizada para estimar o benefício do tratamento no nível do indivíduo, em vez de no nível populacional. A definição e a estimativa de tais parâmetros permitem, se combinadas com *Big Data*, dar suporte ao médico em suas decisões, por enfatizar as melhores escolhas de estratégia terapêutica. Komorowski et al. desenvolveram um modelo computacional capaz de sugerir dinamicamente os melhores tratamentos para pacientes adultos com sepse na UTI.⁽¹³⁾

Limitações atuais e conclusões

Devemos reconhecer a existência de limitações que precisam ser superadas para permitir que o AM direcionado se torne uma realidade no futuro.⁽¹⁴⁾ Primeiramente, enquanto as UTIs hoje geram *gigabites* de dados a cada dia, apenas uma pequena fração deles é acessível atualmente para fins de pesquisa.⁽¹⁵⁾ Em segundo lugar, ainda permanecem questões importantes a respeito de qual a melhor forma de valorizar os *Big Data* médicos e o AM na UTI. São necessários ensaios randomizados e controlados, para demonstrar o benefício da análise preditiva e prescritiva nos pacientes críticos. Entretanto, levando em consideração os avanços recentes, *Big Data* médico e AM oferecem uma oportunidade singular para mudar dramaticamente nosso paradigma, partindo da era da medicina baseada em evidência, na qual as decisões terapêuticas se baseiam essencialmente em evidência oriunda de cortes com base populacional, para uma nova era de suporte médico ideal e personalizado.

REFERÊNCIAS

- Kalbandi I, Anuradha J. A brief introduction on big data 5Vs characteristics and hadoop technology. *Proc Comput Sci*. 2015;48:319-24.
- Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. *Chest*. 2018;154(5):1239-48.
- Mehta N, Pandit A. Concurrence of big data analytics and healthcare: A systematic review. *Int J Med Inform*. 2018;114:57-65.
- Malak JS, Safdari R, Zeraati H, Nayeri FS, Mohammadzadeh N, Farajollah SS. An agent based architecture for high-risk neonate management at neonatal intensive care unit. *Electron Physician*. 2018;10(1):6193-200. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29588819>
- Herasevich V, Pickering BW, Dong Y, Peters SG, Gajic O. Informatics infrastructure for syndrome surveillance, decision support, reporting, and modeling of critical illness. *Mayo Clin Proc*. 2010;85(3):247-54.
- Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035.
- Pirracchio R, Cohen MJ, Malenica I, Cohen J, Chambaz A, Cannesson M, Lee C, Resche-Rigon M, Hubbard A; ACTERREA Research Group. Big data and targeted machine learning in action to assist medical decision in the ICU. *Anaesth Crit Care Pain Med*. 2019;38(4):377-84.
- Dinov ID. Volume and value of big healthcare data. *J Med Stat inform*. 2016;4. pii: 3.
- Gambus P, Shafer SL. Artificial intelligence for everyone. *Anesthesiology*. 2018;128(3):431-3.
- Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med*. 2015;3(1):42-52.
- Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med*. 2015;7(287):287ra71.
- Food and Drug Administration (FDA). U.S. Department of Health and Human Services. Paving the way for personalized medicine: FDA's Role in a New Era of Medical Product Development [Internet]. Maryland: FDA; 2013 [cited 2019 Jan 14]. Available from: <https://www.fda.gov/oc/resources/files/10/10-28-13-Personalized-Medicine.pdf>
- Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med*. 2018;24(11):1716-20.
- Naidus E, Celi LA. Big data in healthcare: are we close to it? *Rev Bras Ter Intensiva*. 2016;28(1):8-10.
- Celi LA, Mark RG, Stone DJ, Montgomery RA. "Big data" in the intensive care unit: Closing the data loop. *Am J Respir Crit Care Med*. 2013;187(11):1157-60.