

Utilização de machine learning em laboratórios de análises clínicas

Machine Learning utilization in Clinical Laboratory

Albert Bacelar de Sousa¹

¹ Médico/ Escola Bahiana de Medicina e Saúde Pública (EBMSP); Coordenador UTI. Salvador, BA, Brasil.

Resumo

Nos últimos anos, as publicações de pesquisa relacionadas ao Machine Learning (ML) aumentaram significativamente na área de patologia e medicina laboratorial. No entanto, apesar dos recentes avanços tecnológicos e do crescente corpo de literatura, existem poucos exemplos de ML implementados na prática clínica de rotina. A pesquisa bibliográfica ocorreu no Medline em 12 de fevereiro de 2020. O foco foi nos artigos publicados desde 2018 (entre 1º de janeiro de 2018 e 31 de janeiro de 2020) para basear a análise em estudos recentes. Concluiu-se que à medida que as ferramentas proliferam, o papel da medicina laboratorial no desenvolvimento, validação e manutenção desses modelos permanece importante, porém pouco definida.

Palavras-chave: Técnicas de Laboratório Clínico; Análise de Dados; Ciência de Dados

Abstract

In recent years, research publications related to Machine Learning (ML) have increased significantly in the field of pathology and laboratory medicine. However, despite recent technological advances and the growing body of literature, there are few examples of ML implemented in routine clinical practice. The literature search took place on Medline on February 12, 2020. The focus was on articles published since 2018 (between January 1, 2018 and January 31, 2020) to base the analysis on recent studies. It was concluded that as the tools proliferate, the role of laboratory medicine in the development, validation and maintenance of these models remains important, however little defined.

Keywords: Artificial Intelligence; Machine Learning; Decision Support Techniques; Neural Networks, Computer

Correspondência

Albert Bacelar de Sousa

E-mail: albert@abacelar.com

Recebido em 16/08/2020

Aprovado em 16/03/2022

DOI: 10.21877/2448-3877.202202058

INTRODUÇÃO

À medida que a demanda por assistência médica de qualidade continua a ascender exponencialmente, o mesmo ocorre com o volume de testes de laboratório. Semelhante a outros setores, as pesquisas no campo da medicina laboratorial começaram a investigar o uso do *machine learning* (ML) para gerir com sucesso o aumento da demanda por serviços e melhorar a qualidade e a segurança.⁽¹⁾

Na última década, o desempenho do ML em tarefas estatísticas de *benchmark* melhorou significativamente devido ao aumento da disponibilidade da computação de alta velocidade em unidades de processamento gráfico, integração de redes neurais convolucionais, otimização do *deep learning* e conjuntos de dados cada vez maiores.⁽²⁾

O consenso atual é que o desempenho geral do ML supervisionado, ou seja, algoritmos que dependem de conjuntos de dados rotulados, atingiu um ponto de inflexão onde os laboratórios clínicos devem procurar aplicativos de missão crítica.⁽¹⁾

Nos últimos anos, as publicações de pesquisa relacionadas ao ML aumentaram significativamente em patologia e medicina laboratorial. No entanto, apesar dos recentes avanços tecnológicos e do crescente corpo de literatura, existem poucos exemplos de ML implementados na prática clínica de rotina. De fato, alguns dos exemplos mais proeminentes de ML na prática atual foram desenvolvidos antes da inflexão recente em publicações relacionadas a ML.⁽³⁾

À medida que a medicina laboratorial continua passando pela digitalização e automação, os laboratórios clínicos provavelmente serão confrontados com os desafios associados à avaliação, implementação e validação de algoritmos de ML, dentro e fora de seus laboratórios. Compreender para que serve o ML, onde pode ser aplicado e o estado da arte e as limitações do campo do ML será útil para a prática de profissionais de laboratório. Este artigo discute as implementações atuais da tecnologia ML nos fluxos de trabalho modernos dos laboratórios clínicos, bem como as possíveis barreiras para alinhar os dois campos historicamente distantes.

Como o ML continua a ser adotado e integrado à complexa infraestrutura de tecnologia de informação em saúde (TIS), a influência do ML na prática de medicina laboratorial permanece uma questão em aberto. Em particular, é importante considerar barreiras à implementação e identificar as partes interessadas para governança, desenvolvimento, validação e manutenção. No entanto, os laboratórios clínicos

devem considerar primeiro o contexto: a aplicação de ML está dentro ou a jusante de um laboratório?

MÉTODOS

O estudo seguiu os itens de relatórios preferenciais para revisões sistemáticas e metanálise (PRISMA).⁽⁴⁾

Identificação de estudos

A pesquisa ocorreu no Medline em 12 de fevereiro de 2020. Realizada uma pesquisa sensível na literatura usando as palavras “machine learning” AND “clinical laboratory” AND “laboratory analysis”. O foco foi nos artigos publicados desde 2018 (entre 1º de janeiro de 2018 e 31 de janeiro de 2020) para basear a análise em estudos recentes.

Seleção de estudos

Todos os resumos foram selecionados, independentemente, por dois revisores; os conflitos foram resolvidos por um terceiro revisor. O texto completo dos resumos selecionados foi avaliado independentemente quanto à elegibilidade por três revisores, e os conflitos foram resolvidos por consenso.

Critérios de inclusão e exclusão

Os estudos eram elegíveis se o artigo descrevesse o desenvolvimento de um modelo de previsão de diagnóstico ou prognóstico para previsão individualizada usando dois ou mais preditores, comparando modelos de previsão com base nos algoritmos de ML; dentro da data dos últimos dois anos (01/18 – 01/2020), ensaios clínicos (*clinical trial*); em inglês, disponível na íntegra e gratuitamente (*full free text*).

Os estudos foram excluídos quando estavam fora da data, entre janeiro de 2018 a janeiro de 2020; modelos desenvolvidos para não humanos, os modelos foram simulados em vez de utilizar participantes, modelos foram desenvolvidos com base em modalidades de dados de alta dimensão; o principal interesse foi avaliar se a modelagem de previsão (ML) está influenciando no setor médico – análises clínicas.

Extração de dados

Focamos em questões metodológicas do desenvolvimento de modelos e aspectos que abordem ML nos exames e diagnósticos laboratoriais. Os itens extraídos incluíram autor, data da publicação, características gerais do estudo, amostra, objetivo, características do modelo ML e resultados do desempenho do modelo.

Análise de dados

Utilizamos análise descritiva para resumir os resultados (Quadro 1). Em cada artigo, identificamos o método de ML. Identificamos várias comparações no mesmo artigo como resultado da implementação de vários algoritmos de

ML, desenvolvendo modelos para mais de um resultado, desenvolvendo modelos baseados em diferentes conjuntos de preditores (por exemplo, uma vez com e uma vez sem medições de laboratório) ou desenvolvendo modelos para vários subgrupos separadamente.

Quadro 1

Análise dos artigos

Autor	Data	Amostra	Objetivo	Modelo ML	Resultados
Bhosale et al. ⁽⁵⁾	2018	Controle 43 Amostra 43	Para avaliar a presença de biomarcadores de proteínas séricas associados às fases iniciais da formação de placas ateroscleróticas carotídeas, foram feitas análises proteômicas quantitativas sem rótulo para amostras de soro coletadas como parte do estudo de risco cardiovascular em jovens finlandeses.	Para identificar o painel de proteínas com o mais alto desempenho discriminativo, a regressão logística penalizada por Lasso, implementada no pacote R glmnet, foi aplicada aos dados proteômicos séricos. Primeiro, todos os candidatos a preditores foram identificados pela redução dos coeficientes de preditores não informativos para zero usando Lasso com validação cruzada de três vezes, repetindo o procedimento de randomização 200 vezes. Em cada dobra, foram consideradas apenas proteínas significativamente diferenciadas em abundância (ROTS; $P < 0,05$).	Com base nisso, um painel de três proteínas, FbIn1c, ApoE e CDH13, foi observado para fornecer a melhor discriminação entre os casos e controles. Com a inclusão de ApoE e CDH13, houve uma melhora estatisticamente significativa no AUROC (0,79, IC 95%: 0,69-0,88, $p = 0,03$). Somente o FbIn1c classificou os casos de controles com um valor de AUROC de 0,67 (IC 95%: 0,56-0,79).
Ko et al. ⁽⁶⁾	2018	1742	Aplicação de algoritmo de ML validado clinicamente para a detecção de doenças residuais com análise por citometria de fluxo multicolor na leucemia mieloide aguda e síndrome mielodisplásica.	Utilizaram duas técnicas de inteligência artificial (IA) para desenvolver um algoritmo de interpretação MFC para detecção de MRD usando uma coorte no mundo real de mais de 1000 pacientes com LBC e MDS com mais de 5000 dados de MFC em amostras de medula óssea.	A alta validade clínica do algoritmo foi demonstrada, através da previsão de resultados bem-sucedidos no cenário pós-indução.
Dobaño et al. ⁽⁷⁾	2019	195	Para avaliar os fatores que afetaram o Ig em cada antígeno, foram ajustados os primeiros modelos de regressão linear multivariáveis e univariáveis (coeficiente, intervalo de confiança de 95% [IC], valores de p).	Foram realizadas análises multimarcas por análise de componentes principais (PCA), matrizes de correlação e análise discriminante de mínimos quadrados parciais de aprendizado de máquina (PLS-DA) usando os pacotes R FactoMineR, Corplot e DiscriMiner, respectivamente.	O estudo forneceu evidências de um efeito positivo de RTS, S nas respostas de anticorpos a certos antígenos que estão associados à proteção. Como os modelos de regressão logística multivariáveis nos quais as IgGs para MSP1 B12, MSP5 e SSP2 (que são aumentados pela vacinação) foram associados à proteção foram ajustados pela vacinação RTS, S, existe um efeito protetor adicional desses anticorpos além da proteção oferecida pela vacina RTS, S.
Waljee et al. ⁽⁸⁾	2019	401	Identificar características dos pacientes com doença de Crohn na linha de base e na semana 8 da terapia com ustekinumab que predizem remissão a longo prazo na terapia com ustekinumab usando proteína C-reativa (PC-R) como biomarcador da atividade da doença; avaliar a vantagem incremental do uso desses preditores em comparação com o nível do medicamento isoladamente.	Usaram a Floresta aleatória (Random forest) que é um método de previsão de aprendizado de máquina que utiliza um conjunto de árvores de decisão para classificar as observações.	A média (DP) do AUROC para o modelo da semana 8 foi de 0,78 (0,042), e o AUROC no conjunto de testes representativo foi de 0,78 (IC 95%, 0,69-0,87). Os preditores mais importantes nesse modelo foram os níveis de PC-R nas semanas 3, 6 e 8 e a proporção sérica de ustekinumab/PC-R nas semanas 3 e 6. Além da PC-R, a variável laboratorial mais importante incluída neste modelo foi a albumina medida na semana 8. A sensibilidade para o modelo da semana 8, com 120 participantes no conjunto de testes, foi de 0,79 e a especificidade, de 0,67.

RESULTADOS

Nossa pesquisa identificou 9 artigos publicados entre 1/2018 e 8/2020; sete dos artigos foram publicados em 2018, dois em 2019 e não foram publicados artigos no ano de 2020.

Com base na leitura dos títulos, foram excluídos cinco artigos, por não terem como tema central dados laboratoriais, ou de resposta clínica laboratorial no ML (Figura 1).

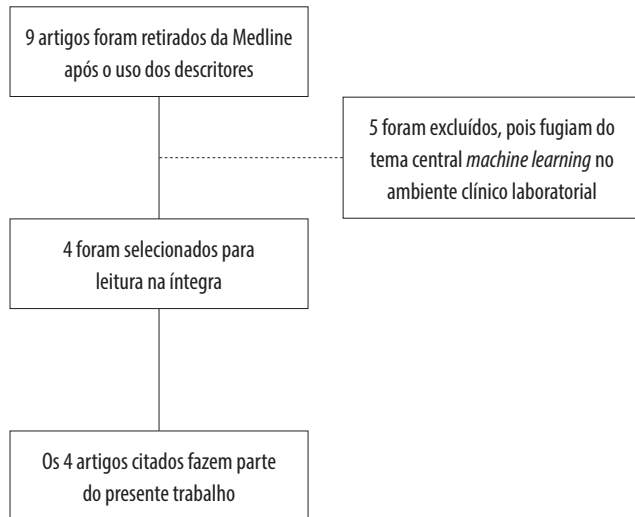


Figura 1

PRISMA flowchart (fluxograma)

DISCUSSÃO

Métodos robustos de ML, como convolução de imagem, redes neurais e aprendizagem profunda (*deep learning*), aceleraram o desempenho do ML baseada em imagem nos últimos anos. As imagens digitais, no entanto, não são tão abundantes em laboratórios clínicos quanto em outras especialidades de diagnóstico, como radiologia ou patologia anatômica, possivelmente limitando futuras aplicações do ML baseadas em imagens na medicina laboratorial.

Os próprios autores desenvolvem os vieses das pesquisas, como Waljee et al.⁽⁸⁾ que consideram o próprio teste imperfeito e assinalam que o contexto clínico deve ser considerado ao se tomar decisões sobre seu uso e interpretação.

Além do número limitado de aplicações comerciais, a pesquisa de ML em medicina laboratorial também tem

aumentado, embora o número total de publicações permaneça relativamente baixo. Nos últimos anos, os pesquisadores investigaram a utilidade do ML para uma ampla gama de conjuntos de dados, como a análise da morfologia dos eritrócitos, morfologia das colônias bacterianas, painéis da tireoide, perfis de esteroides na urina, citometria de fluxo e revisão dos relatórios de resultados dos testes para garantir a qualidade. No estudo de Bhosale et al.,⁽⁵⁾ o fenótipo foi encontrado para assinalar o grupo de risco, e utilizando análise de aprendizado de máquina, foi então identificado um painel de biomarcadores.

Por exemplo, o estudo de Ko et al.⁽⁶⁾ teve uma visão mais positiva e demonstrou que os algoritmos desenvolvidos via IA (inteligência artificial) poderiam realizar a tarefa de classificação em um tempo muito curto (apenas 7 segundos), com cerca de 90% de acurácia na detecção de doença residual em leucemia mieloide aguda e síndrome mielodisplásica. Além disso, os resultados da previsão de resultados no cenário pós-indução demonstraram alto significado prognóstico dos algoritmos de IA.

À medida que os prontuários eletrônicos de pacientes (PEPs) continuam evoluindo e acumulando mais dados, os fornecedores comerciais de PEP procuram expandir seus recursos analíticos e de acesso a dados, começando a oferecer modelos de ML projetados para uso em seus sistemas e, em alguns casos, permitindo o acesso a modelos de terceiros. Os fornecedores geralmente empacotam o *software* de ML no suporte à decisão clínica (CDS), um local cada vez mais popular para misturar ML e medicina clínica – como no caso do artigo de Dobaños et al.,⁽⁷⁾ que procurou ajuda no *machine learning* para medir os níveis de anticorpos por tecnologia quantitativa de conjuntos de suspensões e modelos de regressão aplicados, para analisar fatores que afetam seus níveis e correlatos de proteção.

CONCLUSÕES

Embora algumas instituições tenham integrado com sucesso sistemas de ML locais em seus fluxos de trabalho, poucas fizeram a transição para a prática clínica. Apesar do desenvolvimento de modelos com melhor desempenho, os pesquisadores, por várias razões, costumam encontrar dificuldades com a proverbial última milha da integração clínica. Em particular, a literatura oferece pouca ou nenhuma orientação sobre métricas de desempenho estatístico pelas quais avaliar modelos de ML, o *design* de experimentos de validação clínica

ou sobre como criar modelos de ML mais modulares que se integram às infraestruturas atuais de tecnologia da informação (TI) em medicina de laboratório e fluxos de trabalho.

Com toda a probabilidade, o motivo da lenta adoção de ML pelos laboratórios clínicos, tanto de fontes comerciais quanto de pesquisa, é multifatorial e, sem dúvida, emana de mais do que apenas as limitações intrínsecas da própria tecnologia principal. Semelhante a outras tecnologias que recebem muita atenção, como “big data” ou “blockchain”, o ML continua sendo uma ferramenta que requer uma arquitetura de sistema de suporte. Embora a tecnologia principal esteja demonstrando resultados promissores, é provável que sua prevalência na prática diária permaneça limitada até que desenvolvedores e engenheiros de *software* ofereçam sistemas clínicos de TI que permitam fácil integração com os fluxos de trabalho existentes.

Em seu estado atual, os algoritmos de ML geralmente dependem de dados estruturados para treinamento e subsequente geração de previsões. Embora uma parte significativa dos PEPs contenha dados não estruturados e semiestruturados, as informações de laboratório continuam sendo uma das maiores fontes de dados estruturados, e não é incomum que as ferramentas baseadas em ML confiem fortemente nos dados do laboratório como entrada. À medida que as

ferramentas proliferam, o papel da medicina laboratorial no desenvolvimento, validação e manutenção desses modelos permanece importante, porém pouco definido.

REFERÊNCIAS

1. Tafeit E, Reibnegger G. Artificial neural networks in laboratory medicine and medical outcome prediction. *Clinical chemistry and laboratory medicine*, v. 37, n. 9, p.845-853, 1999.
2. Beam AL; Kohane IS. Big data and machine learning in health care. *JAMA*, v. 319, n. 13, p. 1317-1318, 2018.
3. Torii M, Waghlikar K, Liu H. Using machine learning for concept extraction on clinical documents from multiple data sources. *Journal of the American Medical Informatics Association*, v. 18, n. 5, p.580-587, 2011.
4. Moher D, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, v. 4, n. 1, p.1, 2015.
5. Bhosale SD, et al. Serum proteomic profiling to identify biomarkers of premature carotid atherosclerosis. *Scientific Reports*, v. 8, n. 1, p.1-9, 2018.
6. Ko BS, et al. Clinically validated machine learning algorithm for detecting residual diseases with multicolor flow cytometry analysis in acute myeloid leukemia and myelodysplastic syndrome. *EBioMedicine*, v. 37, p.91-100, 2018.
7. Dobaño C, et al. RTS,S/AS01E immunization increases antibody responses to vaccine-unrelated *Plasmodium falciparum* antigens associated with protection against clinical malaria in African children: a case-control study. *BMC Medicine*, v. 17, n. 1, p.157, 2019.
8. Waljee AK, et al. Development and Validation of Machine Learning Models in Prediction of Remission in Patients with Moderate to Severe Crohn Disease. *JAMA network open*, v. 2, n. 5, p. e193721-e193721, 2019.