



Editorial

Aprendizado de Máquina e Psicometria: Inovações Analíticas na Avaliação Psicológica

Víthor Rosa Franco 

Universidade São Francisco, Campinas-SP, Brasil

Neste editorial, discutirei o que é o aprendizado de máquina (*machine learning*), bem como os seus principais usos na psicometria contemporânea. O aprendizado de máquina pode ser definido como a área de estudos que busca desenvolver algoritmos computacionais que auto-aprimoram por meio da experiência (Mitchell, 1997). Dessa forma, o aprendizado de máquina é considerado como uma subárea da inteligência artificial (IA; Russell & Norvig, 2020). Em modelos tradicionais de IA, todas as relações lógicas entre *inputs* (p.ex., imagens, vetores numéricos, ondas sonoras, entre outros) e *outputs* (p.ex., classificação, resultado analítico, uma mensagem escrita, entre outros) devem ser programadas no código fonte. No entanto, tal estratégia gera rigidez na programação, dado que *inputs* não conhecidos irão gerar falha na geração de *outputs*, o que irá exigir que o código seja sempre modificado. Dessa forma, o advento do aprendizado de máquina possibilitou que relações probabilísticas (ou seja, que consideram a incerteza dos *inputs* e de suas relações com os *outputs*) pudessem ser incluídas em programas de computador e, mais especificamente, IA.

De forma geral, os algoritmos de aprendizagem de máquina são baseados em modelos tradicionalmente usados na análise estatística, como a regressão linear ou a regressão logística (James, Witten, Hastie, & Tibshirani, 2013). No entanto, modelos mais recentes têm implementado procedimentos diversos (p.ex., regressão linear múltipla com regularização de parâmetros; Ghojogh & Crowley, 2019) para gerar flexibilidade e aumentar o desempenho dos algoritmos de aprendizagem de máquina. Devido ao bom desempenho desses novos modelos, muitas vezes, a aprendizagem de máquina é comparada à estatística como se fossem abordagens completamente distintas para lidar com dados (sendo Breiman (2001) a primeira publicação a fazer tal proposta). No entanto, tal afirmação parte do equívoco de equivaler o objetivo dessas áreas. Assim, é correto afirmar que, enquanto a estatística é uma das formas de se analisar a incerteza nos dados, a aprendizagem de máquina usa as análises de incerteza para gerar algoritmos para que um computador aprenda.

Em psicologia, uma discussão similar envolve o uso de modelos estatísticos ou modelos cognitivos/comportamentais (e.g., Farrell & Lewandowsky, 2018; Lee, 2011; Lee & Wagenmakers, 2014). Uma forma de entender a diferença entre esses tipos de modelos é pensar sobre a interpretação dos parâmetros. Em modelos estatísticos convencionais, os parâmetros representam apenas a qual a relação entre variáveis preditoras e critério. Em modelos cognitivos/comportamentais, os parâmetros representam aspectos da cognição derivados de dados comportamentais. Por exemplo, na psicometria, os três parâmetros do modelo logístico de três parâmetros podem ser interpretados como a dificuldade, a discriminação e a probabilidade de chute dos itens (embora haja limitações nessa interpretação; ver Maris & Bechger, 2009). Enquanto isso, nos modelos cognitivos temos modelos estatísticos (ou matemáticos) cujos parâmetros podem, e devem, ser interpretados de acordo com a teoria psicológica subjacente. Ou seja, o parâmetro tem significado teórico, não apenas analítico.

Nesse sentido, modelos e algoritmos de aprendizagem de máquina são similares aos modelos cognitivos/comportamentais: eles descrevem como a unidade de análise (o ser humano ou o computador) deve agir em determinado contexto, levando em conta a incerteza dos dados. Vale salientar que os modelos cognitivos/comportamentais são utilizados para prever o comportamento humano levando em conta o realismo psicobiológico dos processos mentais. Ou seja, um bom modelo cognitivo/comportamental não apenas prediz o comportamento da melhor forma possível, mas também deve respeitar as limitações comportamentais e computacionais impostas pelo corpo humano. Os modelos de aprendizagem de máquina, por outro lado, não precisam respeitar essas limitações, dado que seu objetivo não envolve a necessidade de realismo dos processos de aprendizagem. Em outras palavras, a aprendizagem das máquinas não precisa ser igual à forma com que seres vivos aprendem. Assim, o processo gerador de dados (ou seja, como a aprendizagem em si ocorre) não é tão importante para a aprendizagem de máquinas, sendo a capacidade preditiva dos algoritmos criados a principal medida de desempenho e qualidade.

Aplicar modelos analíticos usados em aprendizagem de máquinas não basta para dizer que se utilizou aprendizagem de máquina em um estudo ou aplicação (apesar do que vemos publicado em diversos artigos, científicos ou de mercado). Por exemplo, utilizar uma rede neural profunda—um tipo de modelo desenvolvido no contexto de aprendizagem de máquina (Nielsen, 2015)—para prever os valores de renda a partir de variáveis sociodemográficas, por si só, não é uma aplicação de aprendizagem de máquina. Pode-se dizer que se utilizou um modelo desenvolvido no contexto de aprendizagem de máquina. Mas, como não há o uso desse modelo para desenvolver um algoritmo para que a máquina aprenda, pela definição de aprendizagem de máquina, não estamos utilizando seus métodos. Por outro lado, caso se desenvolva um algoritmo a partir desse modelo e se use esse algoritmo para, por exemplo, decidir se um determinado indivíduo, ou grupo de indivíduos, deve receber um auxílio do Governo, então sim, houve uma aplicação de aprendizagem de máquina.

Essa discussão inicial é importante para que se possa definir quais as principais aplicações da aprendizagem de máquina em psicometria e em psicologia no geral. Tais aplicações, no entanto, parecem ainda ser escassas (Orrù, Monaro, Conversano, Gemignani, & Sartori, 2020). Não obstante, alguns psicólogos que trabalham com modelagem cognitiva/comportamental foram pioneiros no desenvolvimento contemporâneo da aprendizagem de máquina nos estudos da perspectiva conexionista do cognitivismo, nos estudos de leitura, atenção, categorização e outros, bem como no desenvolvimento de arquiteturas cognitivas e o Processamento de Linguagem Natural (para detalhamento das versões contemporâneas desses estudos, ver Farrell & Lewandowsky, 2018; Ferreira, Sequeira, & Ventura, 2019; Lee & Wagenmakers, 2014; Sun, 2006). Para além desses desenvolvimentos iniciais, neste editorial, interessam-nos três aplicações que, apesar de ainda tímidas, têm diversas ramificações na psicometria: (a) classificação avaliativa; (b) testagem adaptativa computadorizada; e (c) o uso de *big data*.

Em relação à classificação avaliativa, trata-se de prever se um indivíduo pertence a uma determinada categoria (p.ex., se o indivíduo tem ou não um transtorno psicológico) a partir dos resultados de um teste psicológico. Tradicionalmente, ferramentas como Análise Fatorial ou Teoria de Resposta ao Item são utilizadas para avaliar um instrumento e para gerar um escore sem erro de medida, o qual então pode ser utilizado para categorização. Gonzalez (2021) discutiu que usar modelos como regressão logística ou de floresta randômica sem agregar os itens em um escore apenas ou construto podem ter um desempenho tão bom ou melhor do que a abordagem tradicional. Usando um estudo de simulação, o autor demonstrou os contextos em que os modelos de aprendizagem de máquina podem ter desempenho compatível com as análises convencionais de psicometria.

Já a testagem adaptativa computadorizada (*computerized adaptive testing*, CAT; Magis, Yan, & Von Davier, 2017) envolve usar um banco de dados, bem como um algoritmo de decisão, que cria um teste ou avaliação conforme o respondente provê suas respostas. O procedimento mais comum de decisão em um CAT é a medida de informação de Fisher, a qual, com parâmetros dos itens previamente definidos, seleciona o item que maximiza a informação do teste para a estimativa atual do traço latente do respondente. Apesar de esse método ter se demonstrado eficaz em diversos contextos (p.ex., Wainer, 2000), ele é dependente de qual modelo de Teoria de Resposta ao Item foi utilizado para se estimar os parâmetros iniciais dos itens, bem como da quantidade de erro envolvido na estimação desses parâmetros. Com essas limitações em mente, Yan, Lewis e Stocking (2004) propuseram um procedimento de decisão de CAT baseado em árvores de regressão. A abordagem desenvolvida demonstra eficiência mesmo que o modelo de Teoria de Resposta ao Item que gerou os dados não seja conhecido ou até mesmo multidimensional. Os autores (e outros; Ueno & Songmuang, 2010) demonstraram que o procedimento de árvores de regressão pode ter melhor desempenho do que métodos baseados em Teoria de Resposta ao Item paramétrica (para discussões iniciais em procedimentos baseados em Teoria de Resposta ao Item não-paramétrica similares às árvores de regressão, ver Laros & Tellegen, 1991; Huisman & Molenaar, 2001).

A falta de representatividade e, por consequência, a baixa generalização de resultados sempre foi um problema para a psicologia. No entanto, a invenção da *internet*, das redes sociais e dos *smartphones* gerou uma grande revolução (Markowitz, Błaskiewicz, Montag, Switala, & Schlaepfer, 2014). Tais recursos permitem que dados diversos, em grande volume, sejam coletados constantemente, gerando o que se conhece como *big data*. Os *big data*, para além de bancos de dados imensos (ou seja, com milhões ou bilhões de casos e milhares ou milhões de variáveis), abrangem também métodos e procedimentos para organizar e analisar esses dados da forma mais eficiente possível. Consequências diversas, como para a avaliação psicológica, podem ser previstas. Por exemplo, um aplicativo de celular que continuamente coleta dados comportamentais do usuário pode ser capaz de gerar avaliações mais precisas do que os próprios psicólogos e médicos. Isso, tanto em termos de pesquisa como práticos, pode, eventualmente, gerar mudanças profundas à profissão.

Como uma reflexão final, a moral da história deste editorial é que modelos e procedimentos de aprendizagem de máquina podem trazer vantagens à psicometria. No entanto, é importante lembrar que a aprendizagem de máquina visa ao desenvolvimento de algoritmos computacionais para tomada de decisão (ao menos pelo computador). Procedimentos similares ao ajuste de modelos de aprendizagem de máquina, como a modelagem cognitiva/comportamental, trouxeram grandes desenvolvimentos tanto à psicologia como para a aprendizagem de máquina e, portanto,

também merecem um estudo aprofundado. Talvez até mesmo mais aprofundado que a aprendizagem de máquina, dado que seus modelos visam ao maior realismo psicobiológico. De qualquer forma, a aprendizagem de máquina, seja apenas pelos seus modelos quanto pela sua aplicação completa, tem demonstrado usos diversos que, provavelmente, irão impactar nossa concepção de psicométrica e psicologia, seja na academia, ou na sua prática profissional.

Referências

- Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199-231. <https://doi.org/10.1214/ss/1009213726>
- Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge: Cambridge University Press.
- Ferreira, M. I. A., Sequeira, J. S., & Ventura, R. (2019). *Cognitive architectures*. New York: Springer.
- Ghojogh, B., & Crowley, M. (2019). The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. *arXiv preprint arXiv:1905.12787*.
- Gonzalez, O. (2021). Psychometric and machine learning approaches for diagnostic assessment and tests of individual classification. *Psychological Methods*, 26(2), 236-254. <https://doi.org/10.1037/met0000317>
- Huisman, M., & Molenaar, I. W. (2001). Imputation of missing scale data with item response theory. In A. Boomsma, M. A. J., van Duijn, & T. A. B. Snijders (Eds.), *Essays on item response theory* (pp. 221-244). New York: Springer.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York: Springer.
- Laros, J. A., & Tellegen, P. J. (1991). *Construction and validation of the SON-R 5½-17, the Snijders-Oomen non-verbal intelligence test*. Groningen, The Netherlands: Wolters-Noordhoff.
- Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *Journal of Mathematical Psychology*, 55(1), 1-7. <https://doi.org/10.1016/j.jmp.2010.08.013>
- Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge: Cambridge University Press.
- Magis, D., Yan, D., & Von Davier, A. A. (2017). *Computerized adaptive and multistage testing with R: Using packages catr and mstr*. New York: Springer.
- Maris, G., & Bechger, T. (2009). On interpreting the model parameters for the three parameter logistic model. *Measurement*, 7(2), 75-88. <https://doi.org/10.1080/15366360903070385>
- Markowetz, A., Błazkiewicz, K., Montag, C., Switala, C., & Schlaepfer, T. E. (2014). Psycho-informatics: big data shaping modern psychometrics. *Medical Hypotheses*, 82(4), 405-411. <http://dx.doi.org/10.1016/j.mehy.2013.11.030>
- Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.
- Nielsen, M. A. (2015). *Neural networks and deep learning*. San Francisco: Determination Press.
- Orrù, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2020). Machine learning in psychometrics and psychological research. *Frontiers in Psychology*, 10(Article 2970), 1-10. <https://doi.org/10.3389/fpsyg.2019.02970>
- Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach*. New York: Prentice Hall.
- Sun, R. (2006). *Cognition and multi-agent interaction: From cognitive modeling to social simulation*. Cambridge: Cambridge University Press.
- Ueno M., & Songmuang, P. (2010). Computerized adaptive testing based on decision tree. In: 2010 10th IEEE International Conference on *Advanced Learning Technologies, IEEE*, pp. 191-193. <https://doi.org/10.1109/ICALT.2010.58>
- Wainer, H. (2000). *Computerized adaptive testing: A primer*. Mahwah, NJ: Lawrence Erlbaum.
- Yan, D., Lewis, C., & Stocking, M. (2004). Adaptive testing with regression trees in the presence of multidimensionality. *Journal of Educational and Behavioral Statistics*, 29(3), 293-316. <https://doi.org/10.3102%2F10769986029003293>

Como citar este artigo

Franco, V. R. (2021). Aprendizagem de máquina e psicométrica: Inovações analíticas na avaliação psicológica [Editorial]. *Avaliação Psicológica*, 20(3), A-C. <http://dx.doi.org/10.15689/ap.2021.2003.ed>