

EDITORIAL

Ciência de dados e Inteligência Artificial em Medicina

Marco Antonio Gutierrez

Livre-Docente em Informática em Saúde pela Universidade Federal de São Paulo - USP, São Paulo (SP), Brasil

A ciência nasceu há milhares de anos de forma empírica, descrevendo fenômenos naturais. Nos últimos séculos, ela passou a incorporar uma importante componente teórica, utilizando modelos e generalizações. Nas últimas décadas, surgiu uma forte tendência computacional, com a possibilidade de realização de sofisticadas simulações de fenômenos complexos, através da computação científica.

Como **terceiro paradigma**, a simulação de fenômenos complexos, através da computação científica, tem por objetivo criar modelos e métodos matemáticos e computacionais para compreender, analisar e resolver problemas científicos e tecnológicos. Constitui uma grande área de pesquisa de caráter interdisciplinar, fundamentada em conhecimentos científicos e metodologias advindas, fundamentalmente da matemática e da ciência da computação e que busca avançar no desenvolvimento de modelos, métodos, algoritmos e técnicas para simular condições, testar hipóteses e prever a evolução de processos e fenômenos.

É alternativa cada vez mais utilizada para técnicas e observações da ciência experimental, principalmente nos casos em que as medições são impraticáveis, de alto risco ou muito custosas. Encontra aplicações em inúmeras áreas científicas e tecnológicas, pelo que transcende o universo acadêmico e chega ao governo, à indústria, ao comércio, aos serviços e à sociedade.

Na última década, um **quarto paradigma** para a descoberta científica tem surgido devido, principalmente, à disponibilidade de um volume exponencial crescente de dados, oriundos de instrumentos como sequenciadores de genoma, telescópios, fontes de luz, redes de sensores e um crescente número de dispositivos com capacidade de coletar e transmitir dados. Além desse aspecto, fontes de dados, dispositivos de análise e de simulação podem estar interconectados em redes de alta velocidade e de baixo atraso, com capacidade para movimentar grandes volumes de dados.

Entretanto, a geração de grandes volumes de dados por si só não apresenta valor relevante a menos que possam produzir conhecimento. Nesse sentido, o quarto paradigma, que envolve explorar informação em um conjunto massivo de dados, para fomentar descobertas científicas, tem surgido como um complemento essencial aos três paradigmas clássicos: **experimento**, **teoria** e **simulação**.

A complexidade e os desafios para o quarto paradigma envolvem o crescente aumento na velocidade, heterogeneidade e volume na geração de dados de fenômenos físicos e observações.

Nesta nova forma de se fazer ciência, trilhões de bytes de dados podem ser capturados por instrumentos ou gerados via simulação. O acelerador de partículas Large Hadron Collider (LHC) da Organização Europeia para a Pesquisa Nuclear (CERN) captura petabytes (10^{15}) de dados todos os anos, estima-se que nos próximos cinco anos observações e dados de simulações de modelos climáticos atingirão a cifra de hexabytes (10^{18}), o sequenciamento, mapeamento e análise dos vinte e cinco mil genes do genoma humano pode produzir centenas de gigabytes (10^{11}), sendo que o sequenciamento de múltiplos genomas pode produzir centenas de petabytes de dados (10^{17}).

A heterogeneidade desses dados, especialmente em Medicina, sendo em sua maioria não estruturados, podendo variar no espaço-tempo e ainda se apresentarem em formatos complexos, acrescenta outro desafio no esforço de se extrair informação e conhecimento.

É impossível processar essa enorme quantidade de dados sem métodos e processos computacionais avançados. A análise desses dados precisa, obrigatoriamente, utilizar um ferramental determinístico, estatístico e de aprendizado de máquina, que consomem, filtra, manipula, transforma e consolida esses dados, com o objetivo de extrair alguma informação relevante e conhecimento.

A Inteligência Artificial (IA), usando algoritmos capazes de reconhecer um problema, ou uma tarefa a ser realizada, oferece alternativas para a análise dados e para a tomada de decisões, simulando a capacidade humana, mas em uma escala muito superior a do que humanos possam realizar tem se apresentado como um ferramental importante para a tomada de decisões baseadas em dados.

Embora o termo IA tenha sido definido por John McCarthy em 1956⁽¹⁾, a possibilidade das máquinas serem capazes de simular o comportamento humano e, na verdade, pensarem foi proposta muito antes por Alan Turing, que desenvolveu o Teste de Turing para diferenciar humanos de máquinas⁽²⁾.

A incorporação de conceitos de IA em sistemas de apoio à decisão já existem há décadas, porém, foram apenas os aumentos da velocidade de processamento e da capacidade de armazenamento de informações, especialmente com adoção de unidades de processamento gráfico (GPU) para cálculo numérico, que permitiram que grandes volumes de dados fossem analisados, possibilitando a solução de problemas complexos, orientação da tomada de decisões e a realização de tarefas sem receber instruções diretas de humanos.

IA reconhece padrões em imagens, permite interações computadorizadas em linguagem aberta, escrita e falada, percebe relações e nexos, entende conceitos e, não apenas, processa dados, mas pode criar sua própria experiência, também conhecida como aprendizado de máquina.

Em Medicina, além dessas possibilidades anteriores, IA pode também ser utilizada em bases de dados de nascimentos, mortalidade, hospitalizações, doenças de notificação compulsória e em registros de prontuários eletrônicos de pacientes, para indicar a prevalência ou evolução de enfermidades, possibilitando antecipar surtos epidemiológicos e propor medidas preventivas.

O processamento de linguagem natural no registro de dados em prontuários eletrônicos, a adoção de sistemas computacionais na comunicação médico-paciente, o uso de dispositivos vestíveis e corporais para a obtenção de parâmetros fisiológicos e telemedicina visam aperfeiçoar o desempenho do médico no atendimento do paciente e resultarão em importantes transformações na prática médica.

A redefinição da prática médica resultará, necessariamente, em mudanças na formação do médico. Essa preocupação se refletiu no estabelecimento de um consórcio de Universidades pela Associação Americana de Medicina (<https://www.ama-assn.org/amaone/accelerating-change-medical-education>), para discutir mudanças curriculares, ajustando a formação profissional a uma época caracterizada pelo uso intensivo de tecnologias e IA. A mudança da prática médica deverá se refletir em um currículo de formação do médico ajustado a esses novos paradigmas resultantes da mudança de uma época da informação e do conhecimento para uma época caracterizada por novas tecnologias, especialmente IA⁽³⁻⁴⁾.

REFERÊNCIAS

1. Moore J. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazine* 2006;27(4): 87-91;
2. Turing A. Computing machinery and intelligence. *Mind* 1950;49:433-460
3. Wartman SA, Combs CD, Medical Education Must Move from the Information Age to the Age of Artificial Intelligence, *Acad Med* 2018; 93(8):1107-1109. DOI: 10.1097/ACM.0000000000002044.
4. Scott IA, Machine learning and evidence based medicine, *Ann Intern Med* 2018; 169(1):44-46. DOI: 10.7326/M18-0115