

Google cria inteligência artificial para decifrar os segredos do DNA humano

Nova ferramenta analisa como partes ocultas do genoma influenciam genes e podem ajudar a entender a origem de milhares de doenças

Guilherme Levorato

247 - O Google anunciou na quarta-feira (28) o AlphaGenome, uma nova ferramenta de inteligência artificial voltada à análise do genoma humano, com foco em compreender como extensas regiões do DNA influenciam a regulação dos genes dentro das células. A iniciativa busca avançar além da simples leitura do código genético, investigando os mecanismos que controlam quando, onde e como os genes são ativados ou silenciados no organismo, explica a Folha de São Paulo.

Durante a apresentação do AlphaGenome na revista Nature, o vice-presidente de pesquisas da Google DeepMind, Pushmeet Kohli, ressaltou que o sequenciamento completo do genoma humano, concluído em 2003, representou apenas o primeiro passo. “A decodificação do conjunto do genoma humano em 2003 nos deu o livro da vida, mas lê-lo continua sendo um desafio”, afirmou. Segundo ele, apesar de o texto genético estar disponível — formado por cerca de 3 bilhões de pares de nucleotídeos —, compreender sua “gramática” permanece como uma das grandes fronteiras da ciência. “Temos o texto, mas compreender a gramática e a forma como ela governa a vida constitui a próxima grande fronteira da pesquisa”, disse Kohli.

A maior parte do DNA humano não codifica diretamente proteínas. Apenas cerca de 2% das sequências desempenham essa função essencial para o funcionamento dos organismos vivos. Os 98% restantes exercem um papel regulatório complexo, atuando como um maestro que coordena, protege e ajusta a expressão genética em cada célula. É justamente nessas regiões que se concentram inúmeras variantes associadas a doenças, e é esse território que o AlphaGenome se propõe a explorar.

O novo modelo complementa outras ferramentas desenvolvidas pelo laboratório de inteligência artificial do Google, como o AlphaMissense, voltado à análise de

sequências codificantes do DNA, o AlphaProteo, dedicado ao design de proteínas, e o AlphaFold, responsável pela predição de estruturas proteicas e vencedor do Prêmio Nobel de Química em 2024. No caso do AlphaGenome, a inovação está na capacidade de analisar longas sequências de DNA e prever como cada par de nucleotídeos influencia diferentes processos biológicos dentro da célula.

Baseado em técnicas de aprendizagem profunda, o sistema foi treinado com dados provenientes de grandes consórcios públicos, que realizaram medições experimentais em centenas de tipos de células e tecidos humanos e de ratos. Essa base permitiu ao modelo aprender padrões complexos de regulação genética e aplicar esse conhecimento de forma integrada.

Antes do AlphaGenome, já existiam modelos capazes de estudar regiões regulatórias do DNA, mas eles enfrentavam limitações técnicas. Era necessário escolher entre analisar sequências longas, com menor precisão, ou focar em trechos menores, com resolução mais detalhada. Segundo Žiga Avsec, um dos coautores do projeto, compreender plenamente o ambiente regulatório de um gene exige a análise de sequências que podem chegar a até um milhão de pares de nucleotídeos. A nova ferramenta busca superar esse dilema ao combinar extensão e precisão.

Outro diferencial do AlphaGenome é a capacidade de modelar simultaneamente a influência do DNA em 11 processos biológicos distintos. Até então, os pesquisadores precisavam recorrer a diferentes modelos para obter esse tipo de análise integrada. Para Natasha Latysheva, também coautora do estudo publicado na Nature, a ferramenta representa um avanço significativo. “Pode acelerar a nossa compreensão do genoma ao ajudar a cartografar a localização dos elementos funcionais e a determinar seus papéis a nível molecular”, afirmou.

Kohli destacou ainda o caráter colaborativo do projeto. “Esperamos que os pesquisadores o enriqueçam com mais dados e modalidades”, disse. De acordo com o Google, o AlphaGenome já foi testado por cerca de 3.000 cientistas de 160 países e agora está disponível em código aberto para uso em pesquisas não comerciais.

Especialistas externos ao projeto avaliam o modelo de forma positiva, mas com cautela. Ben Lehner, responsável pela área de genômica generativa e sintética no Wellcome Sanger Institute, em Cambridge, afirmou que a ferramenta é muito eficaz, embora ainda tenha limitações. “Identificar com precisão as diferenças em nossos genomas que nos tornam mais ou menos suscetíveis a desenvolver milhares de doenças é um passo crucial rumo a melhores tratamentos”, observou.

Ao mesmo tempo, ponderou que “os modelos de IA são tão bons quanto os dados usados para treiná-los”, ressaltando que muitos conjuntos de dados disponíveis ainda são pequenos e pouco padronizados.

Avaliação semelhante foi feita por Robert Goldstone, responsável pela área de genômica no Francis Crick Institute. Para ele, o AlphaGenome não deve ser visto como uma resposta definitiva para todas as questões da biologia, já que a expressão gênica também depende de fatores ambientais complexos. Ainda assim, considerou a ferramenta essencial para o avanço da área. Segundo Goldstone, ela permitirá aos cientistas “estudar e simular de maneira programática as bases genéticas das doenças complexas”, ampliando as possibilidades de pesquisa e compreensão do funcionamento do genoma humano.

<https://www.brasil247.com/saude/google-cria-inteligencia-artificial-para-decifrar-os-segredos-do-dna-humano>

Veículo: Online -> Site -> Site Brasil 247