



A régua dos algoritmos

Especialistas analisam vantagens e riscos de adotar ferramentas de inteligência artificial para mensurar o desempenho e a contribuição de pesquisadores

FABRÍCIO MARQUES — ilustração ALEXANDRE AFFONSO

Ferramentas de inteligência artificial (IA) poderiam executar tarefas em processos de avaliação científica que hoje são confiadas apenas a revisores humanos, propõe um estudo publicado em novembro no *Journal of Informetrics*. O objetivo do artigo, assinado por pesquisadores da Universidade Federal do Rio Grande do Sul (UFRGS), era identificar critérios e atributos capazes de definir se um pesquisador será premiado com uma Bolsa de Produtividade em Pesquisa (PQ) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Essas bolsas, distribuídas atualmente a cerca de 15 mil pesquisadores, oferecem uma complementação à remuneração que eles recebem em suas instituições como reconhecimento a sua produção destacada e ao trabalho de orientar estudantes.

Foram analisados os currículos Lattes de 133 mil pesquisadores, 14.138 dos quais foram agraciados com as bolsas PQ entre 2005 e 2022. Os autores do estudo utilizaram um conjunto de técnicas de aprendizado de máquina aplicado aos currículos de candidatos a bolsas que conseguiu apontar, com razoável grau de acerto, quais pesquisadores seriam contemplados. A precisão chegou a 80% em uma das categorias de bolsistas, a PQ-2, voltada a pesquisadores mais jovens e que

considera principalmente o número de artigos publicados e de orientações de alunos. “Para os demais níveis, a ferramenta se portou bem, mas com menos acurácia, pois a definição depende de uma análise mais qualitativa”, disse um dos autores do estudo, Denis Borenstein, da Escola de Administração da UFRGS.

Segundo o pesquisador, só foi possível criar o modelo porque há um grande volume de dados sobre a produção dos pesquisadores na Plataforma Lattes, utilizado para treinar os algoritmos de aprendizado de máquina. Na sua avaliação, a ferramenta seria útil ao menos para fazer uma triagem dos candidatos e poupar trabalho dos avaliadores. “Os revisores poderiam analisar um volume menor de propostas de forma mais tranquila e cuidadosa”, afirma Borenstein, um especialista em pesquisa operacional aplicada, abordagem interdisciplinar que utiliza algoritmos e métodos matemáticos e estatísticos para auxiliar em tomada de decisões.

O diretor científico do CNPq, o físico Olival Freire, concorda que a inteligência artificial vai tornar-se útil nos procedimentos de avaliação da agência, mas afirma que sua adoção deve ser criteriosa e paulatina. “Sistemas de IA mal trei-

nados ou mal utilizados podem exibir barbaridades nos resultados. É preciso fazer uma curadoria cuidadosa dos algoritmos para ver se a análise é consistente”, afirma. Freire conta que o CNPq utiliza inteligência artificial em tarefas como a seleção de pareceristas que julgam solicitações de bolsas e projetos de pesquisa. “O sistema roda a lista dos 15 mil bolsistas de produtividade e aponta um pequeno leque de especialistas no tema do projeto. Eles em seguida são contatados pelos técnicos do CNPq”, diz.

Essa estratégia, segundo Freire, previne vieses na escolha dos revisores, como o excesso de convites àqueles que rapidamente aceitam a tarefa ou entregam seus pareceres. O uso de programas de inteligência artificial generativa também passou a ser permitido na formulação de propostas de candidatos a bolsas PQ, desde que eles declarem que utilizaram tais recursos. O diretor do CNPq afirma, contudo, que um uso disseminado da inteligência artificial poderia chocar-se com a abordagem mais qualitativa que o CNPq está buscando dar a seu processo de avaliação. “Alguns comitês de áreas

Análise multidimensional

Engenharia química adota critérios qualitativos de avaliação

A distribuição de Bolsas de Produtividade em Pesquisa (PQ) do CNPq na área de engenharia química está seguindo critérios diferentes dos usualmente adotados por outras disciplinas. Em vez de se limitar aos tradicionais indicadores quantitativos, como número de artigos, citações e alunos orientados, o foco é mensurar a contribuição dos candidatos a bolsistas de forma ampla, avaliando desde o impacto acadêmico de sua produção científica até o seu papel na formação de recursos humanos, o esforço em estabelecer colaborações e a liderança do pesquisador em projetos científicos e de inovação. As regras válidas para o próximo triênio vêm sendo discutidas pelo comitê assessor de engenharia química do CNPq nos últimos quatro anos e começaram a ser implementadas em outubro, depois de serem discutidas

com a comunidade. “A ideia é considerar o caráter qualitativo da pesquisa, de modo a desestimular práticas predatórias de publicação que inflam artificialmente a produção científica”, diz Claudio Dariva, pesquisador da Universidade Tiradentes, em Aracaju, Sergipe, atual coordenador do comitê.

A busca por novos critérios se explica, em parte, pela disputa excessivamente acirrada pelas bolsas de produtividade de engenharia química. Considerando a média das três últimas avaliações para as bolsas PQ (2021, 2022 e 2023), a cada 100 pesquisadores da área que se candidatam, somente 35 têm sucesso, o menor nível de atendimento entre todas as engenharias no CNPq. Para se ter uma ideia da assimetria, a média de atendimento da demanda solicitada para todas as engenharias no período

gira em torno de 45%, com algumas áreas chegando a 60%. “Se usarmos critérios meramente quantitativos para distribuir as bolsas, vários pesquisadores com contribuições importantes à sociedade nem sequer conseguem concorrer. A nossa pergunta norteadora sempre foi: o que é ser um pesquisador produtivo?”, diz Maria Alice Zarur Coelho, pesquisadora da Universidade Federal do Rio de Janeiro (UFRJ), que liderou o comitê assessor durante parte do período da formulação dos novos critérios. “O objetivo é contribuir com uma análise multidimensional do pesquisador, que sirva como uma ferramenta orientadora e privilegie o impacto da pesquisa de forma ampla”, afirma Marisa Beppu, pesquisadora da Universidade Estadual de Campinas (Unicamp), que também liderou o comitê assessor de engenharia química do CNPq.

disciplinares do CNPq estão julgando os pedidos de bolsas de produtividade em duas etapas. A primeira, de caráter mais quantitativo, analisa dados sobre produção científica, citações e número de orientações dos candidatos e pode ter o suporte de algoritmos. Na segunda, os proponentes são convidados a apontar não números, mas suas principais realizações, que podem ser um artigo influente, uma patente ou uma produção artística, e o julgamento é feito por pares. Isso não poderia ser feito de forma apropriada pela inteligência artificial”, explica.

Ferramentas de IA já são usadas para organizar e analisar grandes volumes de dados de pesquisa e até para identificar ou modelar estruturas de proteínas que podem gerar novos medicamentos. O físico Osvaldo Novais de Oliveira Júnior, pesquisador e atual diretor do Instituto de Física de São Carlos da USP, mostrou que a inteligência artificial é capaz de prever, com grande chance de sucesso, se um artigo científico receberá muitas citações. Ele e colegas da USP e da Universidade de Indiana, nos Estados Unidos, publicaram um estudo sobre o assunto no repositório arXiv, que ainda não foi revisado por pares. O trabalho analisou resumos de 40 mil artigos publicados entre 2012 e 2022 na revista *ACS Applied Materials & Interfaces*, publicada pela American Chemical Society, utilizando métodos de inteligência artificial e conseguiu apontar com 80% de precisão quais estavam entre os 20% mais citados – isso, com base apenas nas palavras utilizadas e nos tópicos abordados, sem levar em conta quem eram os autores e as instituições a que pertenciam. Um estudioso de linguística computacional, Novais afirma que o domínio da linguagem humana pelos computadores fará com que eles consigam executar atividades intelectuais de todo tipo e que a gigantesca capacidade de processar dados das máquinas as levará a superar, em breve, a inteligência humana. “É provável que, em um futuro não muito distante, por volta de 2027, atinjamos a singularidade tecnológica, momento em que a inteligência artificial ultrapassará a capacidade humana”, completa.

O diretor científico da FAPESP, o geneticista Marcio de Castro Silva Filho, considera irreversível a tendência do uso de IA no processo de avaliação. “Editoras científicas já utilizam a tecnologia na triagem e na análise de artigos científicos submetidos. As agências de fomento tendem a caminhar nessa direção, desenvolvendo ferramentas que deem suporte para seus assessores”, diz. “Estamos discutindo na FAPESP como desenvolver algoritmos que permitam extrair informações das propostas e auxiliem os avaliadores no momento da revisá-las.” Segundo ele, o

essencial é dar transparência à utilização da inteligência artificial conforme ferramentas desse tipo forem incorporadas, além de deixar claro quais são os critérios que as norteiam.

Mas a viabilidade de utilizar algoritmos em tarefas complexas do processo de avaliação pode demorar, de acordo com a médica Rita Barradas Barata, que entre 2016 e 2018 foi diretora de avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). “Uma coisa é usar algoritmos para processar grandes volumes de dados, outra é interpretá-los de modo a captar todas as nuances necessárias para a análise de projetos.” A pesquisadora, que foi responsável pela conclusão da avaliação quadrienal dos programas de pós-graduação divulgada pela Capes em 2017 (*ver Pesquisa FAPESP nº 260*), diz que dimensões múltiplas do desempenho de cursos de mestrado e doutorado, como o contexto em que operam ou suas vocações regionais, devem ter peso no processo de avaliação e precisarão ser consideradas por algoritmos que vierem a ser desenvolvidos.

Jacques Marcovitch, que foi reitor da USP entre 1997 e 2001, afirma que o uso preditivo de inteligência artificial envolve riscos. Um deles é inibir a adoção dos princípios da chamada “avaliação responsável”, que busca introduzir parâmetros qualitativos, baseados em revisão por pares, na análise de resultados científicos. “Algoritmos têm capacidade de ler uma grande quantidade de conteúdos, mas olham sempre para o passado, para dados acumulados ao longo do tempo. Num era de rupturas, essas ferramentas limitam a identificação e a valorização da ciência que vai moldar o nosso futuro”, afirma. Marcovitch lidera o Projeto Métricas, um esforço que reúne pesquisadores de várias instituições para desenvolver formas abrangentes de mensurar o impacto das universidades na sociedade.

Isso não significa, segundo o pesquisador, que modelos de inteligência artificial não possam ser úteis: o fundamental, ele diz, é que seus resultados sejam usados por pessoas qualificadas, que compreendam suas limitações e saibam interpretá-los. Justin Axel-Berg, também pesquisador do Projeto Métricas, alerta para a falta de transparência dos parâmetros adotados por algoritmos de inteligência artificial generativa. “Seria arriscado usar esses programas para determinar a concessão de recursos públicos para projetos e bolsas. O que responder para um candidato inconformado com o resultado da avaliação? Que foi o algoritmo que disse não?”, indaga. ●