

On Reproducible AI: Towards Reproducible Research, Open Science, and Digital Scholarship in AI Publications.

Odd Erik Gundersen, Yolanda Gil, David W. Aha. 2018

O artigo discute a importância da reprodutibilidade na pesquisa científica e os desafios para alcançá-la, particularmente no campo da IA e aprendizado de máquina. O artigo sugere que a documentação adequada e a adesão a métodos científicos sólidos são cruciais para garantir a reprodutibilidade na pesquisa de IA. A ciência aberta também é enfatizada, o que envolve o compartilhamento de dados, software e outros recursos científicos em repositórios públicos usando licenças permissivas para promover acessibilidade e reutilização. O artigo analisa publicações recentes de IA e destaca as limitações em sua documentação, enfatizando a necessidade de melhores práticas de documentação e disseminação para facilitar a reprodutibilidade, apoiar a ciência aberta e adotar o conhecimento digital. Um estudo de Gundersen e Kjensmo (2018) que analisou o quão bem a pesquisa empírica de IA está documentada para reprodutibilidade. O estudo propõe uma estrutura para reprodutibilidade com três fatores: método, dados e experimento. A documentação é medida usando 16 variáveis sim/não, e três métricas de reprodutibilidade são propostas. O estudo teve como amostra 400 artigos de conferências e as descobertas sugerem que a pesquisa de IA computacional não é documentada sistematicamente o suficiente para apoiar a reprodutibilidade. Poucos artigos mencionam explicitamente o método de pesquisa ou compartilham o conjunto de dados de teste e apenas uma pequena porcentagem compartilha o código-fonte ou especifica sua hipótese e previsão. O estudo tem limitações, mas seus resultados destacam a necessidade de uma melhor documentação na pesquisa empírica de IA. Gundersen e Kjensmo (2018) definem três níveis de reprodutibilidade: R1 (reprodutível pelo experimento), R2 (reprodutível por dados) e R3 (reprodutível pelo método). Para atingir o R1, os pesquisadores devem documentar o método de IA, os dados e a configuração do experimento. Para R2, apenas o método AI e os dados precisam ser documentados, enquanto o R3 requer apenas a documentação do método AI. Se uma equipe independente reproduz os resultados da pesquisa, o nível de documentação fornecida determina a generalidade dos resultados. Pesquisadores independentes confiam mais nos resultados com mais documentação, enquanto o esforço para reproduzir os resultados aumenta com menos documentação. Os pesquisadores originais têm incentivos para documentar com menos vigor para tornar o método mais geral, enquanto os pesquisadores independentes preferem a pesquisa R1. A comunidade de pesquisa, patrocinadores de financiamento, empregadores de pesquisadores e editores devem incentivar e recompensar pesquisas reproduzíveis. O artigo recomenda que os pesquisadores documentem e compartilhem seus dados e código-fonte para que outros possam usá-los e citá-los facilmente. Isso deve fazer parte das práticas diárias de pesquisa, e os comitês de estabilidade devem considerá-lo ao avaliar os candidatos à estabilidade. O artigo sugere resumos estruturados que devem conter a motivação da pesquisa, objetivo, método utilizado, resultados e conclusão, o que dará coerência à narrativa da pesquisa. Essa estrutura ajudará a melhorar a legibilidade e a comparabilidade dos trabalhos de pesquisa.