Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology

Nikhil Agarwal, Alex Moehring, Pranav Rajpurkar, Tobias Salz

Abstract

Artificial intelligence (AI) algorithms have matched or surpassed the performance of human experts in a number of predictive tasks, although human experts have access to contextual information that may not be available for machine predictions. We investigate how best to combine machine predictions with human input in the presence of such contextual information and potential biases in how humans use machine predictions in forming their assessments. Our experiment varies the information environment for radiologists diagnosing patients with potential thoracic pathologies to measure (i) the value of contextual information (clinical history and notes) and (ii) the extent of automation bias/neglect and own-information bias/neglect relative to the Bayesian benchmark. We find that that contextual information improves diagnostic accuracy on average, but providing AI predictions does not always increase accuracy. AI assistance reduces accuracy because subjects partially neglect both their own-information and AI predictions when updating their beliefs. Therefore, AI assistance should be deployed selectively as a function of how predictions are utilized by human experts.