



Edward Shortliffe M.D., Ph.D., Columbia University

The Role of Knowledge in Medical AI: How Smart Can an AI System Be if it Doesn't Know Anything?

Abstract:

The roots and goals of artificial intelligence lie in the study of ideas that enable computers to do the things that make human beings seem intelligent: reasoning symbolically, acquiring and applying knowledge, and manipulating and communicating ideas (Winston textbook of AI, ~1980). The phrase “knowledge is power” then dominated in applying AI in fields such as medicine. In medical AI these notions were viewed as central in areas such as decision support (e.g., “knowledge-based” or “expert” systems), knowledge representation (e.g., production rules, semantic networks, ontologies), machine learning (e.g., knowledge discovery in databases or KDD), and natural language processing. Yet with the evolution of machine learning (ML), and the explosive interest in neural networks and deep learning, some developers are referring to the models or algorithms that emerge from training as “new knowledge”. Should such models, with their impressive analytical power, be viewed as largely mathematical tools for decision support rather than as knowledge? They lack the notions of computational representation that have dominated in the AI field and they offer no insight into how human beings would attempt to solve the same problems. In some fields, including medicine, our human inability to understand or communicate the semantic basis for such ML-based assessments have an impact on a user’s ability to understand and accordingly to trust the results. This accounts for the growing recognition that explainability and interpretability are crucial elements in all clinical applications of AI. Accordingly, powerful ML tools and approaches must be augmented by parallel efforts to seek a merger of knowledge-based methods with the powerful capabilities of data science.