AI for Health - A Primer

Marcel Salathé, Sharada P Mohanty

Digital Epidemiology Lab, EPFL
Campus Biotech, Chemin des Mines 9, 1202 Geneva, Switzerland

Artificial intelligence (AI) - the phenomenon of machines being able to solve problems that require human intelligence - has in the past decade seen an enormous rise of interest. A combination of new machine learning algorithms, increased computational power, and an explosion in the availability of very large data sets ("big data") has led to stunning advances with demonstrations of machines achieving human-level competence at solving clearly defined tasks across many domains. The health sector, one of the most important sectors for societies and economies worldwide, is particularly interesting for AI applications. The potential for AI-assisted health decision making is enormous.

Health is a basic human right. Good health for everyone has for centuries been a key goal of most governments, and public health breakthroughs such as vaccination are generally credited with having saved - and continue to save - billions of lives. In many countries today, the healthcare industry is the largest and / or fastest growing industry, increasingly often spending more than 10% of GDP on health. It is thus not surprising that when a technology reaches new levels of maturity, as is the case with modern AI, the healthcare sector is a key sector for applications of the technology. On the one hand, the potential economic opportunities are big, given the size of the sector; on the other hand, the potential for leveraging new technology for the common good by improving public health can be enormous. It is thus prudent to look at the potential of AI in health by asking a few key questions. First, what is AI, and how does it differ from already existing technologies? Second, what are current applications of AI in health? Third, what is the future potential of AI in health? Fourth, what challenges are there on the journey to unlock this potential?

The term *artificial intelligence* is not new. As an academic research field, it dates back to at least the mid-20th century, and since then has gone through multiple cycles of substantial progress, followed by inflated expectation, followed by disappointment. The current cycle is primarily driven by the extremely impressive progress recently made by deep learning, a branch of machine learning that very efficiently uses artificial neural networks to address harder problems than ever before. Applications of deep learning have achieved human or superhuman performance in many fields such as image recognition and natural language processing. Importantly, the structure of the neural networks, which is ultimately responsible for their performance, is obtained by a process of automated, iterative training. In many cases, no expert-level knowledge is used in the training process, other than direct input and output parameters (e.g. sets of pixels and their associated labels), giving rise to the so-called "end-to-end" learning. In other words, the networks learn to go directly from one end, the input, to the other end, the output, without requiring any domain-specific expertise in between. The resulting network structures are generally very large, with oftentimes billions of parameters, and of such complexity that it is impossible to describe in simple terms how they work, which has led to new challenges concerning their explainability and interpretability.

The fact that computers are increasingly able to interpret images and text as accurately as humans can[1,2], combined with the remarkable computing power in smartphones, open up countless avenues for AI applications in health. Much of the recent work on AI in health has thus gone into applications that revolve around image interpretation, and language understanding. In the field of medical image analysis, one of the most publicized studies was by Esteva et al.[3], demonstrating the accurate classification of skin lesions using a deep neural network that was trained on clinical images, and assessing its performance by comparing its classifications to those made by board-certified dermatologists, revealing the network had reached human accuracy levels. A survey published last year reviewed over 300 papers using deep learning in medical image analysis[4]. The applications of those papers cover brain image analysis, retinal image analysis, chest x-ray and chest CT image analysis, digital pathology image analysis, breast image analysis, cardiac image analysis, abdominal image analysis, musculoskeletal image analysis, fetal image analysis, dermatological image analysis, and others. The tasks the deep learning approach tried to solve generally involved detection, segmentation, or classification. In language understanding, the areas of biomedical text mining, EHR analysis, sentiment analysis on internet-derived data, and medical chatbots have shown promising results[5].

Taking into account that most of the rapid progress has occurred in the past decade, it's interesting to ask what's next. When assessing the potential of AI in health, we need to distinguish between what the technology can in principle do, today, and what will reasonably be technologically feasible in the next decade. Starting with the latter, at the minimum we can expect a steady increase in the performance of the existing models, that is, approaching or surpassing human accuracy in an increasing number of domains. Given the increasing computational power of mobile devices, it is equally easy to expect that the vast majority of the world's population will have direct access to devices that can, in principle, utilize compute-intense AI-powered applications. It's important to note, however, that we already today possess the basic technological capabilities to create such applications in some domains. For example, relatively accurate detection of skin lesions using a state-of-the art mobile phone is technologically feasible. Medical chatbots that can answer basic medical questions are already on the market. AI-powered analyses of radiology scans are increasingly offered to medical institutions, with the approval of the relevant health authorities.

While this progress is exciting, the potential of AI in health also faces the reality of limited data access and cultural, legal and ethical considerations. As most modern AI applications are based on supervised learning - that is, trained on labelled data - access to high-quality labelled data is a major limiting factor. Making such data openly accessible is unlikely to be a general solution to the problem, as health data is often very sensitive data. Trusted data cooperatives with appropriate access management may provide a potential solution. In addition, machine learning approaches must take into account the bias that both text and image-based medical data most likely contain[6]. Further, deep learning models are famously hard to interpret and explain, which may substantially hamper their cultural and practical acceptance, although it's important to remind ourselves that this is not a fundamentally new problem in medicine. As medical decision making increasingly becomes algorithmic (or algorithm-assisted) decision making[7], the practical use of AI in health will benefit from an open and transparent approach in the development and use of these algorithms. Given the speed at which such algorithms can be developed, improved, and deployed, AI has the potential to make first-class medical decision making accessible and affordable to the entire global population.

Marcel Salathé (marcel.salathe@epfl.ch) is a professor at the Schools of Life Sciences, and Information and Communication Sciences, at EPFL, where he leads the Digital Epidemiology Lab. Sharada P. Mohanty (sharada.mohanty@epfl.ch) is a PhD student in the Digital Epidemiology Lab at EPFL.

This short report was written for the the "AI for Good Global Submit", held on May 15-17, 2018, at ITU, Geneva, Switzerland.

References

1. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015. He K, Zhang X, Ren S, Sun J. *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, 1026–1034 (doi :10.1109/ICCV.2015.123)
2. Google's neural machine translation system: Bridging the gap between human and machine translation. 2016. Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, et al. *arXiv*:1609.08144.
3. Dermatologist-level classification of skin cancer with deep neural networks. 2017. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. *Nature* 542, 115–118. (doi:10.1038/nature21056)
4. A survey on deep learning in medical image analysis. 2017. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, et al. *Medical Image Analysis*, 42:60-88
5. Opportunities and obstacles for deep learning in biology and medicine. 2018. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, et al. *J. R. Soc. Interface*; (DOI: 10.1098/rsif.2017.0387)
6. Semantics derived automatically from language corpora contain human-like biases. 2017. Caliskan A, Bryson JJ, Narayanan A. *Science* 6334:183-186: 183-186
7. Dynamic Clinical Algorithms: Digital Technology Can Transform Health Care Decision-Making. 2018. Bell D, Gachuhi N, Assefi N. *Am. J. Trop. Med. Hyg.* 98(1), 9–14 (doi:10.4269/ajtmh.17-0477)