Artificial Intelligence in Healthcare

AAIH ALLIANCE FOR ARTIFICIAL INTELLIGENCE IN HEALTHCARE

A Technical Introduction

Executive Summary

Artificial Intelligence (AI) has been studied by computer scientists for more than 70 years and it is one of the most elusive topics in computer science because of the vast number of techniques employed and the often nebulous goals being pursued. The term 'Artificial Intelligence' itself was coined by John McCarthy in 1956 at the first workshop on the subject at Dartmouth College.[1] But the theory and topics that became known as AI have a much longer history.[2]

Evolution of AI in Healthcare

AI and healthcare have been bound together for over half a century. The DENDRAL project, an early Expert System based on AI techniques from Stanford in the 1960s, aimed at hypothesis formation and discovery in science. The primary focus was to determine organic compound structures by analyzing their mass spectra. A lot of new theoretical and program language work was undertaken to make this possible. It was followed by MYCIN in the 1970s with the goal of identifying infection-causing bacteria and to recommend antibiotics, with dosage adjusted for the patient's weight. The concepts behind MYCIN were then generalized to all internal medicine in the 1980s with the CADUSEUS system, described at the time as the "most knowledge-intensive expert system in existence." Also in the 1980s, several techniques were developed for use in drug discovery. Since then, the number of techniques and uses in healthcare has grown steadily against a wider backdrop of AI "summers and winters" (The summer/winter metaphor has been the comparison of choice for describing the cyclical rise and fall of interest in AI and expectancy/hype around its deliverables). Today we are experiencing unprecedented AI summer that many believe is an integral part of the Fourth Industrial Revolution.

Defining AI

An Intelligent Agent (also known as a Rational Agent) is an autonomous entity that directs its activities toward accomplishing complex goals by making observations of its environment through sensors, processing the inputs, and acting on the environment through actuators (or effectors). Examples of Intelligent Agents are humans, dogs, thermostats, modern airplanes, etc. Intelligent Agents may lack certain elements (such as Software-Only Agents considered separately from underlying hardware) and be different in their degree of autonomy. Following from the definition of Intelligent Agent, AI is the study of artificial intelligent agents and systems, exhibiting the ability to accomplish complex goals. When spoken about today, what is meant by AI is an Intelligent Agent with Machine Learning (ML) at its core. ML may be combined with other algorithms to enable the Agent, such as with *alpha GO*, the Agent that recently beat the world leading GO masters. ML is *the study of algorithms and statistical models that computer systems use to perform specific tasks without using explicit instructions, relying on patterns and inference found in the training data and in the environment*.

Understanding Machine Learning

ML methods can be categorized along a number of different axes: subtype, e.g. supervised or unsupervised; type of model produced, such as, classification, regression, or ranking; area of use, such as, vision, robotics, diagnostics, drug discovery, etc. Supervised ML algorithms take labeled data, such as compounds with biochemical assay measurement labels or images with skin cancer diagnosis labels, and builds a model that can predict these labels for compounds or images without assay measurements or cancer diagnoses. Supervised ML algorithms produce a model that can be used to predict labels on new data and can be distributed independently of the original training data. Unsupervised ML algorithms find patterns in unlabeled data, such as, biomarker or target discovery by determining the biological factors that are important in a population with a disease versus without a disease; or when given a set of compounds produce a set of novel compounds that have the same properties but that are structurally distinct. Unsupervised ML may also produce a model that is independent of the training data, as in the novel compound production examples, or just insights into the dataset, such as, in the novel biomarker or target discovery case. In all uses of ML the training data and the problem being solved for are as central to selecting the right ML algorithm, which results in highest quality models and insights.

1 J. Moor, "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years," *AI Magazine,* vol. 27, no. 4, pp. 87-91, 2006.

2 N. J. Nilsson, The Quest for Artificial Intelligence, Cambridge: Cambridge University Press, 2010.

Understanding the Role of Data

Data are the fuel that the ML engine runs on. As such, the quality, quantity, and composition of the data are critical. The higher quality the data the better the model. The same goes for quantity. But quality and quantity are often competing factors. Lowering the quality standard can often lead to a higher quantity of data. The right choice will be problem dependent and will determine which ML algorithm will produce the superior model or insight. The composition of the training data is arguably the most critical. Particularly in its relationship to the data on which the model or insight will be applied (test data). If the background of training data is different than the background of the test data in the supervised learning case, the resulting model will produce predictions with systematic error, known as Bias. Bias comes from a ML model containing erroneous assumptions. The erroneous assumptions come from the relationship between the training data and the test data, from an inappropriate choice of features to represent the data, or from the machine learning process itself. These aspects when properly orchestrated can actually help to compensate for issues in the other. Bias is most often context specific. Certain biases exist when applied to one test set, but not another. Avoiding all biases is impossible given finite training sets, therefore information about how a ML model was trained, what data it was trained on, what method was used, etc. must accompany the model to inform the application.

Solving the Relevant Problem

The problem being solved for by building a ML model are as important to the model quality as the data. How a model will be used is something that needs to be examined before building a model. This relates to the bias discussion in the last paragraph, but also relates to the question of how the model will be used to support decision making. Will the model be treated as a binary decision tool? If so, what is the tolerance for false positives and false negatives, e.g. how conservative does the model need to be? In the case of skin cancer diagnosis, a high false positive rate will result in unneeded trips to the hospital and a high false negative rate will lead to negative health outcomes. Does the model need to be a regression model? Scientists often want a quantitative output instead of a categorical output, but high quality regression models have a higher data requirements than binary or categorical models. Model building frequency is another consideration. How often do new data arrive? How often will the model need to be retrained: monthly, weekly or continuously? Different modeling approaches and ML algorithms are more amenable to continuous (also known as on-line) or frequent retraining. Finally, how much is known about the underlying process that is being modeled? This is important in helping to determine how to represent the underlying training set to the ML algorithm. ML is concerned with making predictions based on a training set, therefore all correlations of variables in the training set with the labels will be picked up by the ML algorithm. The algorithm will not be able to distinguish between causality and correlation, so to improve the generalizability and trustworthiness of the resultant model, non-relevant correlations need to be removed in the feature sets before model building. This is an issue because in Biology, the underlying causal structure of the process is often not well understood.

The Promise of AI in Healthcare

Intelligent Agents enabled by ML models can be operated at speeds and scales well beyond human capability. ML models have the ability to take in and train on more data than any one person, and ML algorithms can build far more complex models than any human can. Since ML is data-driven, the models can be applied to any problem given proper training data. These Agents will become essential in various Healthcare applications due to the complexity of biology, the rate at which new knowledge is being generated in healthcare, and the reaction speed needed in time-critical decision making. ML is the tool that can enable scientists, clinicians, and medical professionals from biomedical discovery through clinical development to patient care and population health to make better decisions.

By developing and adopting such Agents, we will reduce failure rates and lower drug development costs by increasing the number and quality of available targets, designing and testing fewer molecules that are more effective at treating disease with limited toxicity or adverse events, and selecting the right patients at the right time for the right treatment in clinical trials. In clinical care, ML enabled Agents drive efficiencies in the clinical workflow itself. Further, ML enabled Agents play a significant role in aiding decision-making among health practitioners along the continuum of prevention, diagnosis, treatment and patient follow-up. They permit early, accurate diagnosis by aggregating disparate pieces of information, extracting key patterns from data sets to help identify effective interventions early on when conditions are amenable to treatment. Sample use cases include imaging algorithms to classify cancer and models to determine treatment policies for septic patients. If these successes are extrapolated to the broader health indications and settings, the benefits to society and government will be very significant.

Conclusion

Given the potential benefits of AI in healthcare, but also the real possibility to cause harm, we call for a concerted and collaborative effort to improve industry-wide understanding of the complexities of AI. Critically, the healthcare industry should work together with governments and patients to advance the discussion of responsible AI use. We must work to develop standards that will ensure trustworthiness and transparency in decisions supported by Intelligent Agents as we promote the use of those Agents in all aspects of healthcare to ensure the best possible decisions are always made.