CONFERENCE

Session Report

The Use of Artificial Intelligence and Machine Learning in Clinical Research and Health Care

Speaker
J. Kelly Byram, MS, MBA, ELS
Founder and CEO, Duke City Consulting, LLC, Albuquerque, NM

By Noelle Ochotny, PhD

Artificial intelligence (AI) projects are becoming more common assignments for medical communicators. In fact, a show of hands in this session revealed that approximately one-third of us have worked on a project involving AI and, among those who had not, many anticipated they would in the coming year.

The presentation was organized into 3 sections. The first discusses the basics of AI and machine learning (ML) technology. The second describes how AI tools are developed and implemented. The third identifies and provides examples of the current and emerging applications of AI in clinical research and health care.

The session's focus is on narrow AI, specifically, on a type of ML called supervised learning. It is important to note that not all the AI health applications discussed in this session are implemented in health care yet.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING TERMS

Artificial Intelligence (AI): Leverages computers and machines to simulate the problem-solving and decision-making capabilities of the human mind.

Weak AI, also called **Narrow AI:** This type of AI is limited to a specific task or narrow area.

Strong AI: An AI that has mental capacities and flexible intelligence that mimic the human brain. This type of AI is also sometimes referred to as artificial general intelligence, artificial consciousness, or sentience.

Machine Learning (ML): An AI technique that teaches computers to learn from data.

Algorithm: A set of instructions. In ML, the algorithm learns and evolves without human intervention based on the data it processes. The algorithm builds on commonly used models such as linear regression, logistic regression, Bayesian algorithms, and decision trees. The terms algorithm, model, and tool are sometimes used interchangeably.

Deep Learning: A type of ML that uses an artificial neural network (ANN). ANNs are layers of connected nodes designed to emulate human processing of information.

WHAT ARE SOME CHALLENGES ASSOCIATED WITH AI?

ML model creation requires data and computing resources. Creating the complexity required for a valid ML model can require extensive resources. There are also issues with trust in the ability of the model to make correct decisions that are for the benefit of the patient. Therefore, data and privacy security need to be robust. Having strong data and privacy security can help build trust in the ML model.

CATALYSTS FOR THE DEVELOPMENT OF ML

Recent developments have catalyzed the development of ML models and include

- Big data, which provides the necessary data and resources.
- Cloud storage, which enables organizations to store, access, and maintain large amounts of data required for ML model development. Also consider that the volume of medical data doubles every 8 to 12 months, which requires a lot of storage.
- Powerful computing is essential. ML can be a computationally intensive process, so a powerful computer is needed to handle the load.
- Parallel processing, which allows the ML model to be deployed across multiple processors. This is necessary for the ML algorithm to perform large amounts of computation on large data sets, especially in the deep learning context.
- Maturation of statistics and mathematical methods, which underlie ML.

ISSN 2163-5315

TYPES OF ML

There are 3 common types of ML models, supervised, unsupervised, and reinforcement learning.

Supervised learning uses labeled training data that pairs inputs with outputs (input-output pairs are called examples; a collection of examples is a data set). An application of this type of learning might be an application that predicts if a specific type of tumor is likely to be malignant or benign based on its size, for example. Algorithm training, validation, and testing require a data set be subdivided into a training data set, a validation data set, and a testing data set. Supervised learning occurs via a training loop (Figure).

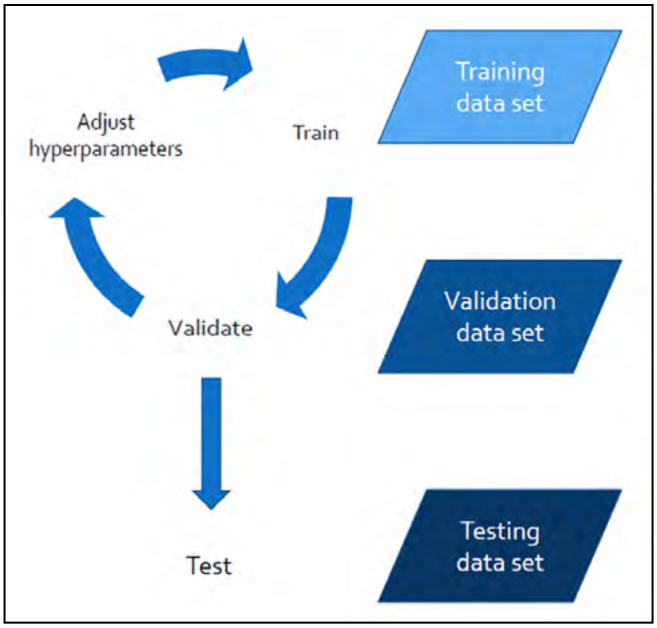

Figure. Supervised learning training loop. Copyright 2022 Duke City Consulting, LLC.

In supervised learning, the percentage of the data set dedicated to each function depends on biostatistical calculations that will inform this decision. This can be similar to the sample size calculations used to determine the sample size in clinical trials. It is important that the data sets are representative of the population so that the resulting ML model is generalizable.

Unsupervised learning uses unlabeled training data—rather, the computer looks for patterns in the data. Some examples of this type of learning are images and pathology data. Used for clustering, segmentation is an example of how this type of learning can be applied.

Reinforcement learning is a reward-and-penalty type of learning beyond the scope of this session.

Deep learning has more complicated models composed of nodes organized in layers. Each layer transforms information and passes it to another layer. The term "deep" refers to the number of layers through which the model operates. In this type of learning, the model can learn from itself to create new features. Ground truth information is fed back into the model to enable the deep learning model to learn from itself. (Ground truthing is a term used in ML that means checking the results of machine learning for accuracy against the real world.) An example of deep learning is facial recognition. Deep learning requires a vast number of resources.

WHAT ARE SOME ML MODELS BEING DEVELOPED IN HEALTH CARE AND CLINICAL RESEARCH?

1. Risk assessment and prevention. The patient completes a questionnaire that includes a personal and family history, and the algorithm can calculate the patient's risk for cancer. To do this, the algorithm uses guidelines such as the National Comprehensive Cancer Network guidelines to determine a patient's risk for certain cancers. The algorithm then goes on to suggest risk reduction strategies and treatment plans for that patient. Is there a role for medical writers? Consider that there are patient-facing and clinician-facing aspects related to the model. Medical communicators may develop information regarding risk reduction, testing, and treatment options that is provided to those audiences. The medical writer can communicate what is going into the model and what is coming out of the model using language specific to the 2 audiences: clinicians and patients. In addition, trust in the model is an ongoing issue. The medical writer plays an important role to develop trust.

Attendees were interested in whether there are guidelines in place on how to communicate the risk assessment and prevention tools to patients in order to obtain informed consent. These tools collect sensitive data from patients and their families, so the patient and the family need to provide informed consent.

The attendees were also interested in whether medical students and residents are receiving training on AI and ML. One attendee reported that their institution, the University of Florida, launched a curriculum on AI development for physicians and clinicians. AI and ML models may affect how medical students and residents receive training. For example, a radiologist has seen thousands of images, and current medical students may not get that experience.

2. Clinical decision support software. There is an ongoing debate about which software is regulated as a device. The United States Food and Drug Administration (FDA) issued a guidance[1] regarding Clinical Decision Support Software to describe the FDA's regulatory approach to Clinical Decision

Support software functions. Consult the guidance for a complete discussion and examples.

Keep in mind that if the software contains ML, then the FDA considers it to be a device. The FDA consistently updates guidelines for AI/ML applications.

3. Diagnosing retinal disease. This is a fast-growing market that includes diabetic retinopathy, a common complication of diabetes. To diagnose retinal disease, a camera takes an image of the patient's retina that is then analyzed using ML. Several papers reported an AI detection rate of retinal disease that is better than the detection rate of clinicians. However, the AI performed worse at diagnosing negative cases.[2,3]

4. Reading and segmenting medical images. Several ML tools are being used in the radiology field. Radiologists who used ML to read medical images worked 65% faster.[4] It is important to note that the use of ML tools improves workflow but does not replace the radiologist.[4] Many of the ML tools listed on the FDA website as being approved are developed for radiology.

5. Predictive modeling. These models help predict outcomes like who will require readmission to the hospital within 30 days of discharge, among others. Predictive modeling can be added to a hospital's electronic health record package. The process to add predictive modeling to an electronic health record package is straightforward because there are vendors who can add the models. For example, the tracking of fall risk, heart failure, and early detection of sepsis can be electronic health record add-ons. However, it is vital that these add-ons meet guidelines requiring model reporting, be useful, fair, and reliable, and are generalizable and transparent. A lack of transparency in an AI model can pose a significant barrier to gaining the trust of patients and clinicians.

6. Drug discovery and development. DeepMind's AlphaFold 2 can predict how a protein folds with an accuracy rate similar to crystallography, but in hours rather than months.[5] AlphaFold 2 radically shortens the identification and development cycles for new drugs, a great boon to biomedical researchers.

7. Nanotechnology. There is hope that this emerging technology field can be applied to cancer diagnostics and cancer therapeutics; however, intratumor and interpatient heterogeneity have posed significant barriers in this area. Application of AI methods to the design and analysis of outcomes have met with some success.

WHAT IS GENERATIVE AI?

Generative AI is a type of ML algorithm that is designed to generate new data based on what it has learned from the data that it has been trained on. This can be used to create new images, text, or other forms of data that mimic the characteristics of the training data. For example, AI can create images of people that look real but who do not exist.[6]

WHAT ARE FUNDERS AND THE FDA LOOKING FOR?

The main concerns of funders and the FDA are found in the Good Machine Learning Practice for Medical Device Development: Guiding Principles document.[7] The concerns that some application developers tend to neglect, in Byram's experience, are included in the following list. Of particular focus is improving the performance of the human-AI team. Another concern is that models degrade over time and need to be retrained.

Key takeaways from this guideline include

1. The multidisciplinary team should work together throughout the total product life cycle to ensure that the AI remains relevant.
2. Make sure that clinical study participants and data sets are representative of the population, so the models developed on them are generalizable.
3. Be sure to emphasize the performance of the human-AI team.
4. Ensure that deployed models have the capacity to be monitored with a focus on improved safety and performance, and appropriate controls are in place to manage retraining risks.

Attendees were curious about the intersection of privacy laws and training data sets. Attendees indicated that, in their experience at their institutions, the patient provides consent for their data to be used in data training sets. To build and maintain trust with the patient, the consent form should include a statement that all patient data will be kept secure, and the data used for ML is deidentified using Protected Health Information guidelines.

CLOSING

Byram closed the presentation by emphasizing the importance of being aware of the FDA and funder guidelines and to consult the references provided in the presentation for further guidance.

Noelle Ochotny is a medical writer at Foremost Medical Communications in Mississauga, Ontario, Canada.

Author declaration and disclosures: *The author notes no commercial associations that may pose a conflict of interest in relation to this article.*

Author contact: *nochotny@fmc.cc*

References

1. US Food and Drug Administration. *Clinical Decision Support Software: Guidance for Industry and Food and Drug Administration Staff.* US Food and Drug administration, 2022. Accessed December 20, 2022. https://www.fda.gov/media/109618/download#:~:text=The%20Food%20and%20Drug%20Administration%20%28FDA%29%20has%20long,therefore%20subject%20to%20FDA%20oversight%20as%20a%20device
2. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-2410.
3. Raumviboonsuk P, Krause J, Chotcomwongse P, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *NPJ Digit Med*. 2019;2:25.
4. Hosny A, Bitterman DS, Guthier CV, et al. Clinical validation of deep learning algorithms for radiotherapy targeting of non-small-cell lung cancer: an observational study. *Lancet Digit Health*. 2022;4(9):3657-3666.
5. AlphaFold. AlphaFold: a solution to a 50-year-old grand challenge in biology. DeepMind. Published November 30, 2020. Accessed October 21, 2022. https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology
6. Hill K, White J. Designed to deceive: do these people look real to you? The New York Times. Published November 21, 2020. Accessed October 21, 2022. https://www.nytimes.com/interactive/2020/11/21/science/artificial-intelligence-fake-people-faces.html
7. Good machine learning practice for medical device development: guiding principles. US Food and Drug Administration. Published October 27, 2021. Accessed October 20, 2022. https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

ADDITIONAL RESOURCES

Adir O, Poley M, Chen G, et al. Integrating artificial intelligence and nanotechnology for precision cancer medicine. *Adv Mater*. 2020;32(13):e1901989.

Artificial intelligence and machine learning in software as a medical device. US Food and Drug Administration. Published September 22, 2021. Accessed October 20, 2022. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

Jaber N. Can artificial intelligence help see cancer in new, and better, ways? National Cancer Institute. Published March 22, 2022. Accessed October 20, 2022. https://www.cancer.gov/news-events/cancer-currents-blog/2022/artificial-intelligence-cancer-imaging

Iwano T, Yoshimura K, Watanabe G, et al. High-performance collective biomarker from liquid biopsy for diagnosis of pancreatic cancer based on mass spectrometry and machine learning. *J Cancer*. 2021;12(24):7477-7487.

Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589.

Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. *Cancer Cell*. 2021;39(7):916-927.

Lu JH, Callahan A, Patel BS, et al. Assessment of adherence to reporting guidelines by commonly used prediction models from a single vendor: a systematic review. *JAMA Netw Open*. 2022;5(8):e2227779.

Roose K. A coming-out party for generative A.I., Silicon Valley's new craze. The New York Times. Published October 21, 2022. Accessed October 21, 2022. https://www.nytimes.com/2022/10/21/technology/generative-ai.html

Stable Diffusion 2.1 Demo. Hugging Face. Accessed October 21, 2022. https://huggingface.co/spaces/stabilityai/stable-diffusion

Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med*. 2021;13(1):152.