Artificial Intelligence In Medicine 129 (2022) 102313

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Machine learning and pre-medical education

Vijaya B. Kolachalama *

Department of Medicine, Boston University School of Medicine, Boston 02118, MA, USA
Department of Computer Science and Faculty of Computing & Data Sciences, Boston University, Boston 02215, MA, USA

ARTICLE INFO

Keywords:
Machine learning
Education
Medical professionals
Pre-medical
Postsecondary education

ABSTRACT

Machine learning and artificial intelligence (AI)-driven technologies are contributing significantly to various facets of medicine and care management. It is likely that the next generation of healthcare professionals will be confronted with a series of innovations that are powered by AI, and they may not have sufficient time during their professional tenure to learn about the underlying machine learning frameworks that are driving these systems. Educating the aspiring clinicians and care providers with the right foundational courses in machine learning as part of postsecondary education will likely transform them as high-tech physicians and care providers of the future.

1. Introduction

The long path to becoming a physician in countries such as the United States begins at least as early as undergraduate/postsecondary education or even in high school. Most medical schools recommend an educational pathway (known as pre-medicine or pre-med) that includes completion of courses focused on the scientific fields of biology, physics, chemistry, organic chemistry, neuroscience, and behavioral sciences along with courses in humanities. As such, pre-medicine is not necessarily a major in most degree-granting institutions. Rather, it is an educational track that undergraduate students pursue prior to matriculation to a medical school. It also involves educational and professional development activities such as volunteer work preferably related to patient care, as well as clinical and research experience, followed by the application to medical schools. A student on a pre-med track may choose an undergraduate major in any field if certain required courses are completed. The pre-med courses are necessary to prepare for the Medical College Admission Test (MCAT) and satisfy most medical school prerequisites that are recommended by the Association of American Medical Colleges (AAMC). Similarly, educational institutes in other countries may have their own set of requirements for postsecondary education and meeting them would position the trainee to pursue a medical degree. While these established pathways are been followed faithfully for several years by accredited universities worldwide, the technological advancements, recently due to the rapid progress in artificial intelligence (AI) are disrupting the practice of medicine and the delivery of healthcare. Indeed, in a 2018 report [1], the World Health Organization noted that digital technologies and AI will be vital tools in achieving their global strategic targets – 1 billion more people benefitting from universal health coverage, 1 billion more people better protected from health emergencies, and 1 billion more people enjoying better health and well-being. These initiatives along with several other national and international efforts to modernize healthcare using AI should serve as motivation to re-design the undergraduate curricula, especially the pre-med pathway and enable the next generation of physicians and other care providers to face the emerging data science revolution.

2. Importance of machine learning education

In April 2018, the United States Food and Drug Administration (FDA) approved the marketing of the first digital health device that used AI to detect diabetic retinopathy (DR) in adults [2]. The software driven by an AI algorithm can process digital images of the patient's retinas and detect the probability of mild or more severe DR. The FDA evaluated data from a study of retinal images obtained from 900 patients with diabetes at 10 primary care sites. In the study, the AI algorithm correctly identified the presence of more than mild DR 87.4% of the time and was able to correctly identify those patients who did not have more than mild DR 89.5% of the time. This software can now be installed within a primary care setting to allow practitioners opportunistic screening of DR during routine patient visits. Since this approval, several other AI-based medical devices and algorithms have undergone regulatory clearance [3], and many more of them are in the pipeline to get reviewed and

* 72 E. Concord Street, Evans 636, Boston 02118, MA, USA.
E-mail address: vkola@bu.edu.

https://doi.org/10.1016/j.artmed.2022.102313
Received 29 October 2021; Received in revised form 23 April 2022; Accepted 27 April 2022
Available online 4 May 2022

potentially gain FDA approval. Evidently, the "end-users" of most of these devices are physicians or other care providers. The question that arises is whether to continue to expect physicians to be plain end-users or if they should get educated on the machine learning algorithms that could potentially facilitate diagnostic decision-making in clinical practice. If the care providers are not necessarily well-versed with computing and data sciences, then there is a high chance that these technologies may lose their value.

Physicians and other care providers also have an important role to play in protecting their patients and other stakeholders from pitfalls and inappropriate application of machine learning technologies. It is important to filter the true value of machine learning from the hype as the care providers need to be aware of the inaccuracies of machine learning, unethical use, and unwanted as well as costlier options for managing patients. For example, a recent study showed that AI systems for detecting COVID-19 using chest radiographs can learn spurious shortcuts even when tested on external datasets, thus pushing forward the need for stronger validation protocols and the use of interpretability techniques for evaluating medical AI systems [4]. On the other hand, studies have also revealed common pitfalls related to the development of medical AI models [5], and potential issues with mandating model explainability as a requirement for clinical deployment [6]. There are ongoing debates on various other topics including model relevance and validity, ethics and trust [7], privacy and security [8], as well as cost and efficiency [9]. Evidence from all these studies should prompt the clinical community to obtain a good understanding in machine learning [10], which can then enable them to be comfortable with the growing use of these technologies in healthcare.

It is worth noting that practitioners in many other fields are perhaps using machine learning-based systems without being well-versed in the underlying computing and machine learning fundamentals. Drawing a parallel with what happens within these non-medical fields, it is unsurprising to expect that care providers routinely use medical imaging equipment or other technologies without a thorough understanding of its fundamentals (physical, electronic, etc.). Nevertheless, the risk of not educating the end-users is that we may see less or inefficient user adoption, resulting in limited or incorrect utilization and reduced benefits of AI in healthcare. As the AI technologies continue to progress, the roles of care providers are likely to change as well. The advantage of educating the care providers is that they can more appropriately utilize the power of these technologies for patient care and management.

Clearly, the timing of when to provide this education matters. Recent articles argued that medical schools need to begin training the next generation of medical professionals by introducing machine learning courses within their curriculum [11–17]. There are similar perspectives to encourage machine learning education when physicians are going through advanced clinical training and beyond [18–20]. It is important to realize that change happens slowly, especially when medical schools strive to follow Abraham Flexner's *uniformly arduous and expensive* brand of medical education [21]. An alternative and potentially attractive possibility is to teach machine learning principles and its applications to undergraduate students as they prepare for medical school, at least in those institutions that offer such programs.

3. Pre-medical track provides an opportune window for machine learning education

Introduction to machine learning as part of postsecondary education for students on the pre-med track has several advantages. First, students are in their early formative years, and hence they are likely to gain broad insights related to principles of machine learning and their applications. Since calculus and statistics courses are already mandatory in most schools, students will have the opportunity to take additional courses on these topics along with courses in computing and programming before signing up for a machine learning course. As such, data science and machine learning courses meld calculus and statistics into a closed loop, reinforcing their value and significance. Most universities have already begun to design new sets of interdisciplinary courses related to teaching data science to undergraduate students that build on learning outcomes [22]. Such modules or degree programs are likely to provide students with exciting career paths that one could consider along with medical school such as MD/PhD, MD/MBA, MD/MPH or MD/JD programs, with specific emphasis on data science and machine learning. As such, machine learning skills are in high demand and even trainees who may not end up pursuing medical education are likely to benefit from these courses, both in their careers and personal life. Second, most medical school curricula are already cramped with several foundational and systems-based courses within their first two years to accommodate core clinical experiences in the third and fourth years. Moreover, assessment in undergraduate medical education, which drives much of learning, is largely focused on preparation for licensing exams. The tight schedule and exam-driven focus leave limited scope for the medical education offices to design a comprehensive course or a module that would allow students to gain some understanding of machine learning and its applications. These practical limitations can be overcome if the matriculant joins the medical school with at least a rudimentary knowledge on a few topics related to machine learning.

4. Curricular recommendations

The pre-med curriculum in the accredited universities should consider accommodating at least one foundational course on machine learning and another one focused on the application of machine learning in healthcare and medicine. As part of the foundational course, students can gain preliminary knowledge on the capabilities and limitations of the principal methodologies of data-driven, model-based prediction and decision-making, including inferential statistics, data mining, and machine learning. As part of their application-oriented course, students can develop the skills necessary to assemble computational pipelines and deliver reproducible data analysis of structured and unstructured biomedical datasets. Students can also develop the ability to assess the societal impacts of data-centered methods, including adherence to policy, privacy, security, and ethical norms. It is imperative that students acquire foundational knowledge in calculus, statistics, and computer programming prior to embarking on the machine learning courses. The spectrum of required exposure to machine learning is broad and can be tailored based on the interests and skills of the trainees. While some trainees will go on to become researchers that will develop AI systems (requiring in-depth knowledge of data science), other will be consumers of AI applications (requiring more general knowledge). For those that wish to fully engage in developing AI systems, even these two courses may not suffice and therefore pursuit of an undergraduate degree in computer science, mathematics and engineering may be encouraged.

Graduates from the program will be ready to contribute to the art, science, and engineering of the data-driven processes that are woven into all aspects of society, economy, and public discourse. They will be ready to pursue a healthcare career in which they contribute to the synthesis of knowledge through methodical, generalizable, and scalable extraction of insights from data, as well as to the design of new information systems and products that enable actionable use of those insights toward discovery, clinical diagnosis, patient management and innovation in a wide range of medical applications.

We are still in the early stages of appreciating the full impact of data science in medicine. Thus, adding machine learning courses to the long list of AAMC-recommended or another organization-based pre-med courses might sound impractical for universities to implement curricular changes, and challenging for students to take all of them within the stipulated time. The purpose of this article is therefore not to enforce sweeping changes in the pre-med curriculum but suggest a simple recommendation to increase flexibility on the course selection. For example, all the required pre-med courses related to biology, chemistry, humanities, organic chemistry, and physics can be made mandatory for

one semester, leaving another semester open to taking other courses. In this framework, students will gain more breadth of knowledge based on the existing pre-med courses along with courses related to machine learning and its applications. The accredited universities that already have or those that are currently planning to design undergraduate programs focused on data science are well-placed to offer machine learning courses that can be integrated within the pre-med track. Note that this article mainly recommends pre-med curricular changes in the United States and Canadian universities. Indeed, educational institutions in Australia, Ireland, and South Korea provide a choice to obtain a pre-medical degree [23]. Therefore, educational systems within such countries have an option to propose machine learning courses during postsecondary education.

5. Conclusion

There is growing evidence that AI frameworks driven by machine learning algorithms have the potential to accelerate the workflow of clinicians and other care providers. The future of utilizing AI to improve healthcare is exciting and we are likely to witness more use cases that address routine and challenging clinical applications. To fully realize the promise and evaluate the pitfalls of AI, we need to consider how the next generation of trainees who are aspiring to join the healthcare workforce would need to understand the tenets of machine learning and its medical applications. Medical schools as well as residency and fellowship programs should continue to find ways to offer machine learning training modules despite their overly tight schedules. A forward-thinking initiative will be to offer introductory machine learning courses as part of pre-medical education at accredited institutions. This article can be viewed as a call to the community to pursue these recommendations.

Declaration of competing interest

None.

Acknowledgments

V.B.K. is a founding member of the Faculty of Computing & Data Sciences at Boston University. He teaches an introductory machine learning course within the Graduate Medical Sciences program at Boston University School of Medicine, and acknowledges support provided by Boston University's Broadening Experiences in Scientific Training (BEST) program. V.B.K. also acknowledges support from the Karen Toffler Charitable Trust, the American Heart Association (20SFRN35460031), and the National Institutes of Health (R01-HL159620 & R21-CA253498).

References

[1] Big data and artificial intelligence for achieving universal health coverage: an international consultation on ethics. WHO/HMM/IER/REK/2018.2; 2018.
[2] FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye; 2018.
[3] Benjamens S, Dhunnoo P, Mesko B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. NPJ Digit Med 2020; 3:118. https://doi.org/10.1038/s41746-020-00324-0.
[4] DeGrave AJ, Janizek JD, Lee SI. AI for radiographic COVID-19 detection selects shortcuts over signal. medRxiv; 2020. https://doi.org/10.1101/2020.09.13.20193565.
[5] Roberts M, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell 2021;3:199–217. https://doi.org/10.1038/s42256-021-00307-0.
[6] Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health 2021;3. https://doi.org/10.1016/s2589-7500(21)00208-9. e745-e750.
[7] Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS Med 2018;15. https://doi.org/10.1371/journal.pmed.1002689.
[8] Price II WN, Cohen IG. Privacy in the age of medical big data. Nat Med 2019;25: 37–43. https://doi.org/10.1038/s41591-018-0272-7.
[9] Pianykh OS, et al. Improving healthcare operations management with machine learning. Nat Mach Intell 2020;2:266–73. https://doi.org/10.1038/s42256-020-0176-3.
[10] Aslam TM, Hoyle DC. Translating the machine: skills that human clinicians must develop in the era of artificial intelligence. Ophthalmol Therapy 2021;11:69–80. https://doi.org/10.1007/s40123-021-00430-6.
[11] Kolachalama VB, Garg PS. Machine learning and medical education. NPJ Digit Med 2018;1:54. https://doi.org/10.1038/s41746-018-0061-1.
[12] James CA, Wheelock KM, Woolliscroft JO. Machine learning: the next paradigm shift in medical education. Acad Med 2021. https://doi.org/10.1097/ACM.0000000000003943.
[13] McCoy LG, et al. What do medical students actually need to know about artificial intelligence? NPJ Digit Med 2020;3:86. https://doi.org/10.1038/s41746-020-0294-7.
[14] Pucchio A, Eisenhauer EA, Moraes FY. Medical students need artificial intelligence and machine learning training. Nat Biotechnol 2021;39:388–9. https://doi.org/10.1038/s41587-021-00846-2.
[15] Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ 2019;5: e16048. https://doi.org/10.2196/16048.
[16] James CA, Wheelock KM, Woolliscroft JO. Machine learning: the next paradigm shift in medical education. Acad Med 2021;96:954–7. https://doi.org/10.1097/ACM.0000000000003943.
[17] Blease C, et al. Machine learning in medical education: a survey of the experiences and opinions of medical students in Ireland. BMJ Health Care Inform 2022;29. https://doi.org/10.1136/bmjhci-2021-100480.
[18] Slanetz PJ, Daye D, Chen PH, Salkowski LR. Artificial intelligence and machine learning in radiology education is ready for prime time. J Am Coll Radiol 2020;17: 1705–7. https://doi.org/10.1016/j.jacr.2020.04.022.
[19] Tajmir SH, Alkasab TK. Toward augmented radiologists: changes in radiology education in the era of machine learning and artificial intelligence. Acad Radiol 2018;25:747–50. https://doi.org/10.1016/j.acra.2018.03.007.
[20] Wood MJ, Tenenholtz NA, Geis JR, Michalski MH, Andriole KP. The need for a machine learning curriculum for radiologists. J Am Coll Radiol 2019;16:740–2. https://doi.org/10.1016/j.jacr.2018.10.008.
[21] Flexner A. Medical education in the United States and Canada. From the Carnegie Foundation for the advancement of teaching, bulletin number four, 1910. Bull World Health Organ 2002;80:594–602.
[22] De Veaux RD, et al. Curriculum guidelines for undergraduate programs in data science. Ann Rev Stat Appl 2017;4:15–30. https://doi.org/10.1146/annurev-statistics-060116-053930.
[23] Orr R, Jain A. The case for shortening medical education. Niskanen Center; 2020.