Datathons in Medical Education: A Case Study and Best Practice Recommendations

Michael Steven Yao, Clara Sun, Steve J Stephen, Lathan Liou

Submitted to: JMIR Medical Education
on: June 24, 2024

Table of Contents

Datathons in Medical Education: A Case Study and Best Practice Recommendations

Michael Steven Yao[1, 2, 3] BS; Clara Sun[1, 4] BS; Steve J Stephen[1, 5, 6] MBA; Lathan Liou[1, 7] MPhil

[1]MDplus New York US
[2]Perelman School of Medicine University of Pennsylvania Philadelphia US
[3]Department of Bioengineering University of Pennsylvania Philadelphia US
[4]School of Medicine Case Western Reserve University Cleveland US
[5]School of Medicine and Dentistry University of Rochester Rochester US
[6]Simon Business School University of Rochester Rochester US
[7]Icahn School of Medicine at Mount Sinai New York US

Corresponding Author:
Lathan Liou MPhil
Icahn School of Medicine at Mount Sinai
1 Gustave L Levy Pl
New York
US

Abstract

Background: As artificial intelligence and machine learning become increasingly influential in clinical practice, it is critical for future physicians to understand how such novel technologies will impact the delivery of patient care.

Objective: We describe a trainee-led, multi-institutional datathon as an effective means of teaching key data science and machine learning skills to medical trainees. We offer key insights on the practical implementation of such datathons and analyze experiences gained and lessons learned for future datathon iterations.

Methods: We detail a recent datathon organized by MDplus, a national student-run nonprofit consisting of over 3,000 medical trainee members and physician-innovators. To assess the efficacy of the datathon as an educational experience, a short opt-in post-datathon survey was sent to all registered participants. Survey responses were de-identified and anonymized before downstream analysis to assess the quality of datathon experiences and areas for future work.

Results: Our virtual datathon was attended by approximately 200 medical trainees across the United States. A diverse array of medical specialty interests were represented amongst participants, with 44% of survey participants expressing an interest in Internal Medicine, 33% in Surgery, and 19% in Radiology. Participant skills in leveraging Python and R for analyzing medical datasets improved after the datathon, and survey respondents enjoyed participating in the datathon (average score: 4.23 / 5).

Conclusions: The datathon proved to be an effective and cost-effective means of providing medical trainees the opportunity to collaborate on data-driven projects in healthcare. Participants agreed that the datathon improved their ability to generate clinically meaningful insights from data. Our results suggest that datathons can serve as valuable and effective educational experiences for medical trainees to become better skilled in leveraging data science and artificial intelligence for patient care.

(JMIR Preprints 24/06/2024:63602)
DOI: https://doi.org/10.2196/preprints.63602

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Datathons in Medical Education:

A Case Study and Best Practice Recommendations

Michael S Yao, B.S.[1,2,3], Clara Sun, B.S.[1,4], Steve J Stephen, MBA[1,5,6], Lathan Liou, MPhil[1,7]

[1] MDplus, New York, NY 10029

[2] Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

[3] Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104

[4] School of Medicine, Case Western Reserve University, Cleveland, OH 44106

[5] School of Medicine and Dentistry, University of Rochester, Rochester, NY 14642

[6] Simon Business School, University of Rochester, Rochester, NY 14642

[7] Icahn School of Medicine at Mount Sinai, New York, NY 10029

Corresponding Author:

Lathan Liou, MPhil

Icahn School of Medicine at Mount Sinai

1 Gustave L Levy Pl

New York, NY 10029

Tel: +1(201)398-3507

Email: lathan.liou@icahn.mssm.edu

Running Title: Datathons in Medical Education: A Case Study and Best Practices

Abstract

Background As artificial intelligence and machine learning become increasingly influential in clinical practice, it is critical for future physicians to understand how such novel technologies will impact the delivery of patient care.

Objective We describe a trainee-led, multi-institutional datathon as an effective means of teaching key data science and machine learning skills to medical trainees. We offer key insights on the practical implementation of such datathons and analyze experiences gained and lessons learned for future datathon iterations.

Materials and Methods We detail a recent datathon organized by MDplus, a national student-run nonprofit consisting of over 3,000 medical trainee members and physician-innovators. To assess the efficacy of the datathon as an educational experience, a short opt-in post-datathon survey was sent to all registered participants. Survey responses were de-identified and anonymized before downstream analysis to assess the quality of datathon experiences and areas for future work.

Results Our virtual datathon was attended by approximately 200 medical trainees across the United States. A diverse array of medical specialty interests were represented amongst participants, with 44% of survey participants expressing an interest in Internal Medicine, 33% in Surgery, and 19% in Radiology. Participant skills in leveraging Python and R for analyzing medical datasets improved after the datathon, and survey respondents enjoyed participating in the datathon (average score: 4.23 / 5).

Conclusion The datathon proved to be an effective and cost-effective means of providing medical trainees the opportunity to collaborate on data-driven projects in healthcare. Participants agreed that the datathon improved their ability to generate clinically meaningful insights from data. Our results suggest that datathons can serve as valuable and effective educational experiences for medical trainees to become better skilled in leveraging data science and artificial intelligence for patient care.

Introduction

The exploration of machine learning (ML), artificial intelligence (AI), and other data science (DS)-driven technologies is becoming increasingly popular within clinical medicine. Given the rapidly growing presence of ML in healthcare innovation, it is important for both current and future physicians to understand the fundamentals of ML technology and how they may help inform clinical decision-making.

However, DS and AI education in current medical school curricula is lacking. Despite recent efforts to integrate AI learning objectives into medical education [1], few US medical schools have formally integrated AI-based topics into their curricula. Pupic et. al. [2] and Civaner et. al. [3] report studies of small self-selected groups of medical students and residents participating in both student- and faculty- led electives covering the fundamental theory behind AI applications for medicine. However, opportunities facilitating real-world experience remain limited.

One potential method for hands-on AI education popular across many fields of science and engineering is the “hackathon” and related “datathon”, which are short competitions in which teams of students work together to create new solutions to domain-specific challenges through leveraging real-world data and algorithms. Oyetade et al. [4] offer a scoping review of hackathons and found that such events help students learn both technical and soft skills, and argue that hackathon-based pedagogies be incorporated in classroom environments. Silver et al. [5] describe a hackathon event for current attendings in clinical practice, and found that study participants were better equipped to accelerate specialty-focused innovation after the hackathon. However, similar hackathon events specifically designed for medical students and other undergraduate trainees is not well described in the literature.

In this article, we describe a datathon hosted by MDplus, a 501(c)3 national student-run organization dedicated to empowering the next generation of physician-innovators. We describe the structure of the event, present data on educational outcomes, and offer resources and recommendations for putting together similar events in the future. Our results suggest that datathons and similar events may be an effective means for AI education for medical students.

Methods

Timeline and Participant Recruitment

The 2023 MDplus Datathon (herein referred as “the datathon”) ran a total of 25 days virtually (**Fig. 1**), and was organized by the medical student-led executive team of MDplus consisting of a core datathon planning team of 8 medical students. MDplus’ Slack community, monthly newsletter, and social media pages (LinkedIn, Instagram, Twitter) were used to advertise the datathon. Over the span of three months prior to the start of the datathon, the organizing team recruited sponsors, mentors, knowledge experts, and judges through the MDplus and personal networks. Registration for the event was limited to current medical students, residents, and graduate students in the United States. To provide a fair learning environment for trainees, attending physicians were not eligible to participate. Participants were asked to form their own teams of three to five individuals.

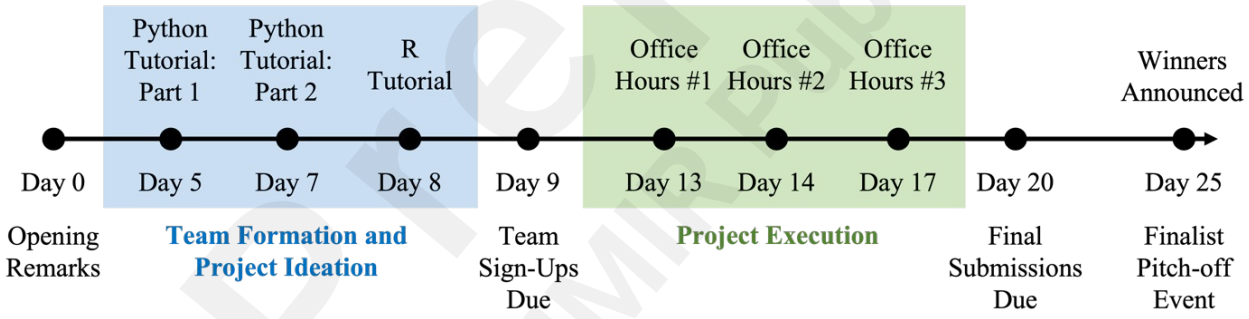

Figure 1: Overview of the Datathon Event. The MDplus datathon ran for approximately 4 weeks and was loosely divided into two parts: (1) Team Formation & Project Ideation and (2) Project Execution.

Dataset

The theme of this datathon was value-based care (VBC), which refers to a healthcare delivery model in which providers are held accountable for improving patient outcomes. In a VBC system, providers are often rewarded with incentivized payments based on quality of care, provider

performance, and the patient experience [7]. Utilizing the Medical Information Mart for Intensive Care (MIMIC)-IV dataset [6], participating teams were tasked with thinking critically about quantitative methods, conducting appropriate analyses (e.g. visualization, statistics, and other computational tools), and contextualizing clinical insights into actionable proposals that solved a problem related to VBC for relevant stakeholders. Briefly, the MIMIC-IV dataset contains anonymized patient data aggregated from over 500,000 patients at the Beth Israel Deaconess Medical Center between 2008 and 2019. Variables from this rich data set include EKGs, medical imaging studies, health records, and patient lab values and outcomes among others. All participants in the datathon were required to sign a data use agreement and complete responsible data handling training in order to gain access to the MIMIC-IV dataset.

Resources and Support

An official datathon page [8] was created for participants as a central hub with instructions, registration, and materials for the event. Links to the datathon's Github Repository were provided with written tutorials and example code including: 1) Downloading and Overview of the MIMIC-IV Dataset; 2) Introduction to Python; and 3) Introduction to R. Optional workshops and private Zoom events with experienced data scientists were offered to students, including Python and R bootcamps, oral presentation workshops, and a pre-recorded Zoom talk with physicians in the VBC space. Communication and announcements throughout the datathon were conducted through Slack.

Submission Format and Judging Criteria

Teams were asked to submit a written technical report of their work and to record a five minute-long oral presentation highlighting key contributions and findings. Participants were free to use any programming language or software to perform their analysis. Judging criteria included statistical rigor, relevance to the datathon theme (VBC), creativity of visualization and/or analysis, and team diversity.

Final Showcase

An internal set of 4 blinded judges composed of members of the MDplus datathon organizing committee evaluated the initial anonymized submissions and selected 7 finalist teams to present at the final datathon showcase. Each team played their recorded 5-minute oral presentations and were allotted 2 minutes immediately after for responding to judge questions. A panel of five judges–recruited for their diverse range of expertises in the VBC space–evaluated the finalists' submissions. Three of the judges are healthcare executives, four are practicing clinicians, and one is a product manager.

Post-Datathon Survey

Upon the conclusion of the datathon, an anonymous 16-question survey was electronically sent to all registered participants. The optional survey asked respondents questions pertaining to team demographics, medical education status, medical specialty interest, familiarity with technical and computational tools, and subjective datathon quality. Participants were asked to rate their familiarity with quantitative tools before and after the datathon on a four-point scale (1 = no familiarity, 4 = a lot of familiarity). To assess the efficacy of the aforementioned technical Python and R tutorials for datathon participants, we compared against participant subjective familiarity with quantitative tools —namely, GitHub and Microsoft Excel—that were not taught explicitly as a part of the datathon. Data were analyzed using Fisher's Exact test in Python 3. To better characterize participant experiences during the datathon, survey respondents also rated their agreement with a set of five standardized statements regarding (1) overall enjoyment of the datathon; (2) VBC topic understanding; (3) ability to identify problems in healthcare; (4) ability to generate insights from data; and (5) likelihood of future datathon participation. Participant sentiment was quantified using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) [9].

Results

Datathon Logistics

The virtual datathon kick-off was attended by 200 medical students, residents, and graduate students across the United States. Twenty-eight teams (109 participants) registered for the datathon and 13 (46%) of the initial registered teams submitted a final project. Among the submitted projects, seven were chosen as finalists to present at the synchronous virtual showcase. The seven projects addressed a variety of topics related to VBC, including chronic kidney disease underdiagnosis, the efficacy of social work referrals, and readmission rates for alcohol-related conditions among others. The final showcase was followed by the announcement of the three winning projects, marking the conclusion of the datathon event.

The organization-accrued cost of organizing and running the datathon was $46 per participant after accounting for expenses such as prize money, computing resources for participants, technical skill-based workshops, and other resources that were provided during the datathon.

Survey Results

Out of the 109 registered participants, 43 (39%) completed the post-datathon survey (**Table 1**). A majority who completed the survey identified as male (74%, *n* = 32) and were under the age of 25 (65%, *n* = 28). Survey respondents self-reported as Asian (65%, *n* = 28), White (23%, *n* = 10), Middle Eastern or North African (5%, *n* = 2), and Hispanic or Latinx (2%, *n* = 1), and 5% (*n* = 2) preferred not to say. 97% of survey respondents were medical students (**Table 2**). There was a wide range of medical specialty interests amongst the 36 medical trainee survey respondents, with Internal Medicine (*n* = 16), Surgery (*n* = 12), and Radiology (*n* = 7) being the most popular specialties.

Age	**Count (%)**
Age < 25	28 (65%)
25 ≤ Age < 30	13 (30%)
30 ≤ Age < 35	1 (2%)
35 ≤ Age	1 (2%)
Self-Reported Ethnicity	**Count (%)**

Asian	28 (65%)
Hispanic or Latinx	1 (2%)
Middle Eastern or North African	2 (5%)
White	10 (23%)
Prefer not to say	2 (5%)
Gender	**Count (%)**
Male	32 (74%)
Female	11 (26%)
Sexual Orientation	**Count (%)**
Heterosexual or straight	42 (98%)
Bisexual, gay, lesbian, or other	1 (2%)
Disability Status	**Count (%)**
Does not identify as a person with a disability	40 (93%)
Does identify as a person with a disability	3 (7%)
Current Education Status	**Count (%)**
Medical student or resident physician	36 (84%)
Other	7 (16%)

Table 1: Demographic information of *N* = 43 post-datathon survey participants.

Familiarity with quantitative tools, Python, R, Github/Gitlab, and Microsoft Excel, before and after participating in the datathon was assessed (**Table 2**). As expected, participant familiarity with GitHub/Gitlab and Microsoft Excel did not significantly change before and after the datathon (GitHub/Gitlab: *P*=.98; Microsoft Excel: *P*=1.00; Pairwise Fisher's Exact Test). In contrast, subjective participant familiarity with R and Python showed greater evidence of improvement after the datahon (R: *P*=.89; Python: *P*=.38; Pairwise Fisher's Exact Test). While the improvement in R- and Python-associated technical skills did not reach the traditional threshold for statistical significance, likely due to the limited sample size of the study, our analysis nonetheless supports the notion that targeted educational tutorials during the datathon event can empower participants with improved technical skills relevant to data science applications in medicine.

Current Medical Education Status	**Count (%)**
First year medical student	11 (31%)
Second year medical student	14 (39%)
Third year medical student	5 (14%)
Year-out medical student	3 (8%)
Fourth year medical student	2 (6%)

Resident physician	1 (3%)
Medical Specialty Interests	**Count**
Anesthesia/Critical Care	6
Cardiology	1
Dermatology	4
Emergency Medicine (EM)	3
Family Medicine (FM)	1
Internal Medicine	16
Mental Health Counseling/Therapy	2
Neurology	5
Obstetrics and Gynecology (OB/GYN)	3
Ophthalmology	5
Pediatrics	4
Physical Medicine and Rehabilitation (PM&R)	1
Plastic Surgery	1
Psychiatry	5
Radiology	7
Surgery (General or Unspecified)	12
Orthopedic Surgery	2
Not currently exploring a medical specialty	1

Familiarity	**Python**		**R**		**GitHub/GitLab**		**Microsoft Excel**	
	Before	After	Before	After	Before	After	Before	After
None	5 (12%)	2 (5%)	11 (26%)	8 (19%)	9 (21%)	8 (19%)	2 (5%)	2 (5%)
A little	7 (16%)	11 (26%)	10 (23%)	12 (28%)	7 (16%)	7 (16%)	6 (14%)	6 (14%)
Some	11 (26%)	7 (16%)	12 (28%)	13 (30%)	16 (37%)	14 (33%)	12 (28%)	13 (30%)
A lot	20 (47%)	23 (53%)	10 (23%)	10 (23%)	11 (26%)	14 (33%)	23 (53%)	22 (51%)

Table 2: Datathon Participatnt Analysis. (Top) Current medical education status and medical specialty interest information for *N* = 36 post-datathon survey participants filtered by medical student and resident physician status. (Bottom) Familiarity with common technical and computational tools before and after datathon participation from a total of *N* = 43 survey responses. Values are reported as count (%).

Figure 2 examines the participant experience quantified by participant agreement with a set of standardized statements. Overall, the survey respondents enjoyed participating in the datathon (mean ± standard deviation Likert score: 4.23 ± 0.60). Using the Likert scale between 1 to 5, participants affirmed that the datathon improved their understanding of the VBC topic (4.00 ± 1.03), their ability to identify problems in healthcare (4.05 ± 0.86) and also generate meaningful insights

from data (4.07 ± 0.85). Participants also agreed on their interest in participating in similar datathon/ hackathon events in the future (4.14 ± 0.95).

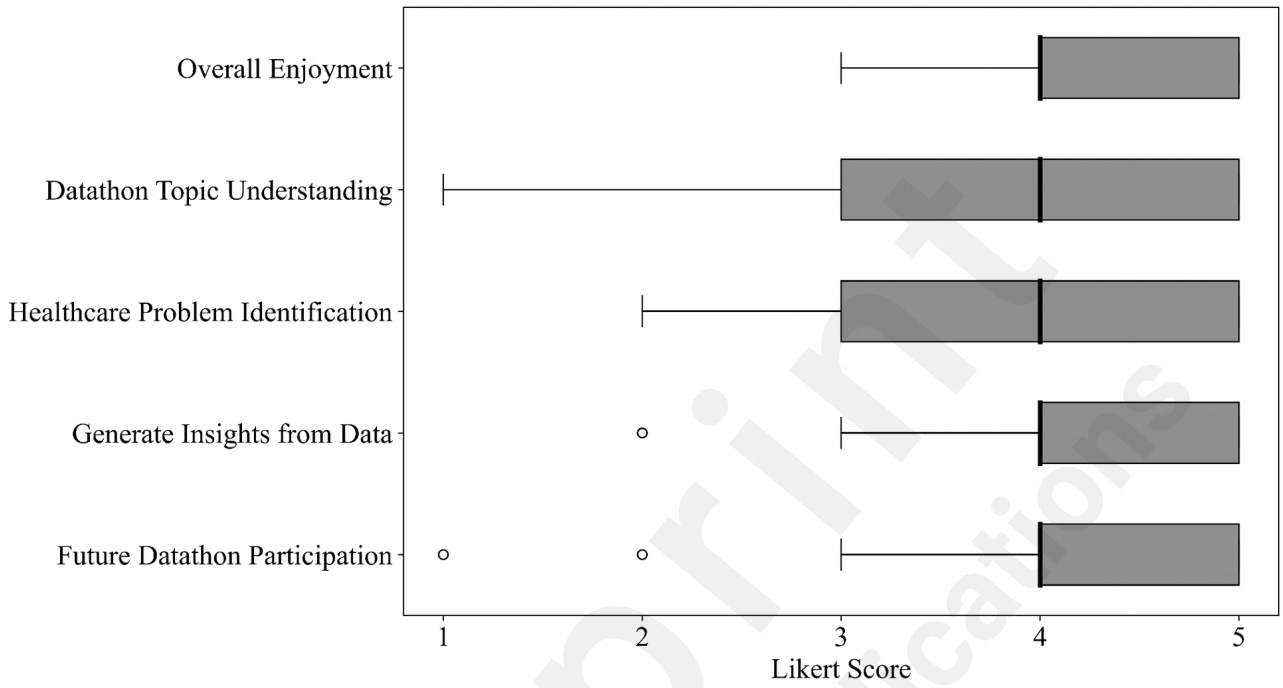

Figure 2: Box plot visualizing survey results assessing for subjective datathon quality. Overall Enjoyment was assessed by the survey question “I enjoyed participating in the datathon.” Datathon Topic Understanding was assessed by the survey question “The datathon improved my understanding of value-based care (VBC).” Healthcare Problem Identification was assessed by the survey question “The datathon improved my ability to identify problems in healthcare.” Generating Insights from Data was assessed by the survey question “The datathon improved my ability to generate clinically meaningful insights from data.” Future Datathon Participation was assessed by the survey question “I intend to participate in other datathon/hackathon events in the future.” Participant scores correspond to the following: (1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree.

A majority of respondents agreed that the datathon improved their understanding of VBC, their ability to identify problems in healthcare, and their ability to generate clinically meaningful insights

from data.

The survey also included an open-ended response option for participants to provide any additional comments. There was a mix of short, positive comments and comments that offer suggestions for future events. Based on our qualitative analysis, key areas for improvement to consider for future datathon iterations include (1) ensuring a balanced distribution a technical skills between participating teams; (2) expediting the team creation process; and (3) offering additional technical workshops and tutorials to participants. Representative example unedited participant comments are shown below:

> *I think in the future, it'd be more effective to make sure each team at least has a "senior" tech lead (someone with 3-5+ years of tech experience) and a "junior" tech lead (1-2 years) to ensure there is great education for all parties involved, as well as greater quality of work. This is of course for folks seeking out teams and not those who already have a team formed that they are comfortable with.*
>
> *…I feel like the team creation process could've been a little faster and I was only able to join a team around halfway into the datathon which didn't give us enough time to work on our idea. But overall, I really appreciate the effort and time put in by everyone involved and I definitely hope to be involved in this again!*

Discussion and Conclusion

To our knowledge, our national, student-led datathon is the first to bring together teams of medical students, residents, and graduate students to propose data-driven solutions within value-based care. In just under a month, 109 registered participants collectively proposed 13 data-driven insights to improve patient care using real-world data.

There are also limitations associated with our study. Due to resource constraints, the number of participants in the datathon limited the power of our statistical analysis. As a result, many of experimental results did not meet the conventional cutoff for statistical significance. Nonetheless, many of our findings and trends noted in our preliminary work provide initial evidence that datathons

can act as effective means for teaching real-world skills to medical trainees in using machine learning and data science for patient care improvement. Finally, our datathon was coordinated virtually with participants joining from many areas of the United States. While there are both benefits and drawbacks to a virtual datathon (as opposed to their in-person counterparts), we leave a rigorous comparison between their utilities in modern medical education paradigms for future work.

Ultimately, the goal of this datathon was to provide opportunities for students, especially medical students, to improve their data skills and to identify data-driven solutions to problems in healthcare. Participants practiced using hands-on data science and artificial intelligence to explore meaningful clinical problems and voiced a collective interest to continue participating in similar initiatives in the future. Overall, our results and collective experiences suggest that datathons can be valuable within undergraduate medical education.

Acknowledgments

The authors thank the teams at Hugging Face, Doximity, Merck, and Conduce Health for their generous support and sponsorship of the 2023 MDplus datathon event, and the datathon judges Reza Alavi, MD, MHS, MBA; Caroline Berchuck, MD, MPH; Amit Phull, MD; Sid Salvi; and Kathryn Teng, MD, MBA, FACP for lending their time and expertise in helping run a successful datathon event. The authors also thank Eric Shan, Katie Link, Julia Bondar, Vamsi Chodisetty, and other members of the MDplus community and executive team for their help in organizing and running the datathon event.

References

1. Sun L, Yin C, Xu Q, Zhao W. Artificial intelligence for healthcare and medical education: A systematic review. Am J Transl Res 15(7):4820-28. (2023). PMID: 37560249
2. Pupic N, Ghaffari-Zadeh A, Hu R, Singla, Darras K, Karwowska A, Forster BB. An evidence-based approach to artificial intelligence education for medical students: A systematic review. PLOS Digit Health 2(11):e0000255. (2023). https://doi.org/10.1371/journal.pdig.0000255
3. Civaner MM, Uncu Y, Bulut F, Chalil EG, Tatli A. Artificial intelligence in medical education: A cross-sectional needs assessment. BMC Med Educ 22(772). (2022). https://doi.org/10.1186/s12909-022-03852-3
4. Oyetade K, Zuva T, Harmse A. Educational benefits of hackathon: A systematic literature review. World J on Educational Technology Current Issues 14(6):1668-84. (2022). https://doi.org/10.18844/wjet.v14i6.7131
5. Silver JK, Binder DS, Zubcevik N, Zafonte RD. Healthcare Hackathons Provide Educational and Innovation Opportunities: A Case Study and Best Practice Recommendations. J Med Syst 40(7):177. (2016). https://doi.org/10.1007/s10916-016-0532-3
6. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, Pollard TJ, Hao S, Moody B, Gow B, Lehman LH, Celi LA, Mark RG. MIMIC-IV, a freely accessible electronic health record dataset. Scientific Data 10(1). (2023). https://doi.org/10.1038/s41597-022-01899-x
7. Teisberg E, Wallact S, O'Hara S. Defining and implementing value-based health care: A strategic framework. Acad Med 95(5): 682-5. (2020). https://doi.org/10.1097/ACM.0000000000003122
8. MDplus. MDplus DS & AI - Datathon. Accessed 18 Jan 2024. URL: http://ai.mdplus.community/datathon/2023
9. Likert R. A technique for the measurement of attitudes. Arch Psychol 140: 1-55. (1932).