Machine Learning for Health (ML4H) 2021

Subhrajit Roy* subhrajitroy@google.com
Google Research

Stephen Pfohl* spfohl@stanford.edu
Stanford University

Girmaw Abebe Tadesse girmaw.abebe.tadesse@ibm.com
IBM Research

Luis Oala luis.oala@hhi.fraunhofer.de
Fraunhofer Heinrich Hertz Institute (HHI)

Fabian Falck‡ fabian.falck@stats.ox.ac.uk
University of Oxford

Yuyin Zhou‡ yzhou103@stanford.edu
Stanford Univeristy

Liyue Shen liyues@stanford.edu
Stanford Univeristy

Ghada Zamzmi alzamzmiga@nih.gov
National Institute of Health (NIH)

Purity Mugambi pmugambi@umass.edu
University of Massachusetts Amherst

Ayah Zirikly azirikly@jhu.edu
Johns Hopkins University

Matthew B. A. McDermott† mmd@mit.edu
Massachusetts Institute of Technology

Emily Alsentzer† emilya@mit.edu
Massachusetts Institute of Technology, Harvard Medical School

List of organizers: Ben Kompa (Finance Chair), Dani Kiyasseh (Communications Chair), Emma Rocheteau (Publications Chair), Fabian Falck (Workflow Chair), Ioana Bica (Outreach Chair), Stephen Pfohl (Program Chair), Subhrajit Roy (Program Chair), Taylor Killian (Virtualization Chair), Yuyin Zhou (Workflow Chair), Alexander Jenkins (Outreach Submission Mentorship Sub-Chair), Ayah Zirikly (Communications Sub-Chair), Elena Sizikova (Virtualization Website Sub-Chair), Elham Dolatabadi (Program Logistics Sub-Chair), Ghada Zamzmi (Outreach Submission Mentorship Sub-Chair), Girmaw Abebe Tadesse (Program Speaker Sub-Chair), Jessica Schrouff (Outreach Career Mentorship Sub-Chair), Liyue Shen (Workflow Submission System Sub-Chair), Luis Oala (Program Thematic Session Sub-Chair), Meg Sharma (Virtualization Gather.Town Sub-Chair), Purity Mugambi (Outreach Reviewer Mentorship Sub-Chair), Scott Fleming (Virtualization Gather.Town Sub-Chair), Young Joon (Fred) Kwon (Virtualization Website Sub-Chair), Emily Alsentzer (General Chair), Matthew McDermott (General Chair).

1. Introduction

The first Machine Learning for Health (ML4H) symposium[§] was held virtually on December 4, 2021. In response to the growing ML4H community, for the first time, ML4H took place as a stand-alone event rather than a workshop at the Neural Information Processing Systems (NeurIPS) conference. The proceedings of ML4H contains eighteen accepted papers.

The symposium invited submissions comprising machine learning research on relevant problems in health and biomedicine.

ML4H 2021 featured two submission tracks: a proceedings track, which encompassed full-length submissions of technically mature and rigorous work, and an extended abstract track, that would accept less mature, but innovative research for discussion. Accepted publications of both types were given a platform for presentation, whether through an oral or poster presentation. The goal was to provide a venue to publish high-quality work, while still enabling the lively discussions that have made the ML4H workshops worthwhile in the past.

In this front matter, we provide an overview of the ML4H 2021 symposium, including its mentorship programs, the paper selection process, and the submission statistics (Section 2). In Section 3 we analyze the accepted works, and offer commentary on trends in research observed in this field, building on analyses of the ML4H workshops from 2019-20 (Dalca et al., 2020; Sarkar et al., 2020). In Section 4, we comment on the composition of the ML4H community. Finally, we close with acknowledgments, including a list of organizers and reviewers for ML4H 2021.

[*] Equal contribution.
[†] Equal contribution.
[‡] Equal contribution.
[§] http://ml4h.cc

2. Symposium

ML4H 2021 symposium was held virtually on December 4, 2021. In accordance with the virtual format, keynotes and spotlight talks were pre-recorded using SlidesLive, panels were held live using Zoom, and poster and roundtable sessions were hosted in Gather.

2.1. Program

ML4H 2021 featured eight invited talks across academia and industry, six spotlight talks from authors of papers accepted at the venue, a best paper, and a best thematic paper award talk. The latter originated from the fact that ML4H hosted a thematic session on real-world robustness and generalization in machine learning for health and meta-reviewers were asked to nominate papers for this award.

The program also included panel discussions between the invited speakers. Speakers at the event were Yoshua Bengio, Shalmali Joshi, Xiao Liu, Sendhil Mullainathan, Rory Sayres, Uri Shalit, Karandeep Singh, and Aisha Walcott-Bryant. The program included poster sessions for accepted papers and extended abstracts. Moreover, ML4H hosted seven roundtable discussions to foster discussion between the participants and senior researchers in the field on several topics of high relevance to the ML4H community. The topics of the roundtable sessions were:

1. *Causality and inductive bias for stability, robustness, and generalization & Detecting failure modes of machine learning systems.* The senior chairs of this roundtable were Federico Cabitza, Lukas Ruff, Wojciech Samek, and Adarsh Subbaswamy. The junior chairs were Utkarshani Jaimini, Kaushik Manjunatha, and Md. Golam Rasul.

2. *What are important clinical problems to solve? A brainstorming session with clinicians.* The senior chairs of this roundtable were Jesse Ehrenfeld, Naomi Lee, Vidur Mahajan, Mark Sendak, and Sergio Uribe. The junior chairs were Kaivalya Deshpande and Huiqi Yvonne Lu.

3. *Multimodal learning in healthcare and representation learning on clinical data.* For this roundtable we had Marzyeh Ghassemi and Polina Golland as senior chairs, and Sneha Jha and Rishab Khincha as junior chairs.

4. *Population health.* The senior chairs were David Buckeridge, Rumi Chunara, and the junior chairs were Serifat Folorunso, Favour Nerrise, and Hang Yuan.

5. Fairness and ethical AI in healthcare. The senior chairs for this roundtable were Roxana Daneshjou, Judy Gichoya, Jeff Lockart, Rohil Malpani, and Andreas Reis. The junior chairs were Jerry Fadugba, Elora Schörverth, and Girmaw Abebe Tadesse.

6. *Regulation in Health AI.* For this roundtable we had Shada Alsalamah, Thomas J. Fuchs, and Regina Geierhofer as senior chairs, and Payal Chandak and Rutwik Shah as junior chairs.

7. *Medical conversations.* For this roundtable we had Sandeep Konam and Zachory Lipton as senior chairs.

2.2. Paper Selection

Submission Statistics. Although ML4H 2021 was held virtually similar to 2020, it continued to feature strong interest from the ML4H community, with 123 total submissions. The total number of submissions decreased from 202 in 2020 and 309 in 2019 which is potentially due to the fact that ML4H was held as a symposium for the first time as opposed to a NeurIPS workshop like in previous years. The program committee consisted of 26 meta-reviewers and 298 reviewers who completed 695 total reviews. At least four reviews were conducted for each proceedings track and extended abstract track submission. Each submission also received a meta-review. The organizers felt that the meta-reviewers reduced the variance in review quality and led to more consistent acceptance decisions. Meta-reviewers reported during the meta-reviewers sessions that they were satisfied with the review process.

Out of the 59 papers submitted to the proceedings track, 18 were accepted into the proceedings (30.50% acceptance rate). The review form for papers submitted to the proceedings track included a question for the reviewer in regards to their recommendation to transfer the paper from the proceedings track to the extended abstract track. Six submitted full papers were asked to transfer to the extended abstract track. Out of the 64 papers submitted to the extended abstract track, 26 were accepted (40.62% acceptance rate). As a result, there were 32 extended abstracts in total, after including the papers transferred from the proceedings track. The extended abstracts were given the opportunity to be included in an ML4H arXiv index at https://arxiv.org/abs/2112.00179 Falck et al. (2021).

2.3. Mentorship Programs

Submission Mentorship Program
In an effort to improve submission quality and foster both current and future collaboration, ML4H hosted a submission mentorship program with two stages. The first stage was 2.5 months long and was for mentees who would like to get feedback on the direction

of the design, experiments, and results for a work they plan to submit to ML4H 2021. In this stage, 30 mentees intending to submit were paired with 18 mentors based on mutual research interests. The second stage lasted 1.5 months and allowed mentees to get feedback on a complete draft, and was suitable for people who were interested in receiving advice from an experienced mentor to enhance the quality of a paper that they plan to submit. In this stage, our team paired 20 mentees with 12 mentors.

Overall, participants reported that the effort was a success. More than 85% had a good overall experience with the Submission Mentorship Program, and 86% of the mentors and mentees expressed interest in participating in the program again in the future years. Moreover, 79% of the participants said that the program had a moderate or major impact on the final submission. However, mentors and mentees both identified time constraints as the biggest roadblock preventing a successful mentor-mentee experience.

Reviewer Mentorship Program

The ML4H symposium also included a Reviewer Mentorship Program whose purpose was to train junior reviewers, foster new connections and relationships in the ML4H community, and ultimately improve the quality of the review process.

The core of the mentorship program was a feedback session where senior reviewer mentors provided feedback to a reviewer mentee on their reviews. Given the short reviewing timeline, mentees submitted drafts of their reviews to their mentor over email one week before the review deadline and feedback sessions occurred the following week, and mentors were free to structure the feedback session how they preferred. It was expected that mentors read their mentees' assigned papers and reviews and formulated their feedback prior to the feedback session with their mentee. However, the role of the mentor was to provide feedback to ensure that reviews were high-quality, constructive, and fair, rather than to serve as an additional reviewer. Overall, 40 mentors, 89 mentees participated. While some mentees didn't reach out to their mentors (hence didn't participate in the program), overall mentors were happy with the mentees they were matched with and expressed interest in collaborating in future.

Career Mentorship Program

The Career Mentorship Program took place on the day of ML4H 2021. The program focused on pairing mentees with mentors that could provide advice on career-related topics, including developing a long-term research plan, doing healthcare research in industry, and work-life balance. Less experienced attendees were encouraged to sign up as mentees and more senior community members were encouraged to sign up as mentors. Mentors were matched with up to three mentees based on their time availability and selected topics of interest and expertise. Each group of matched mentors and mentees was allocated a specific time slot and location in the Gather.Town space.

3. Analysis of Works

3.1. Structured Data Analysis

Alongside their submissions, authors were asked to answer structured questions about their methods and focuses. Figure 1 shows the frequencies of author-reported topics for papers accepted to ML4H. Authors were allowed to select multiple topics for their papers and the top five categories were Medical Image Analysis/Computer Vision, Interpretability, Pretraining/Transfer Learning, Electronic Health Records, and Representation Learning.

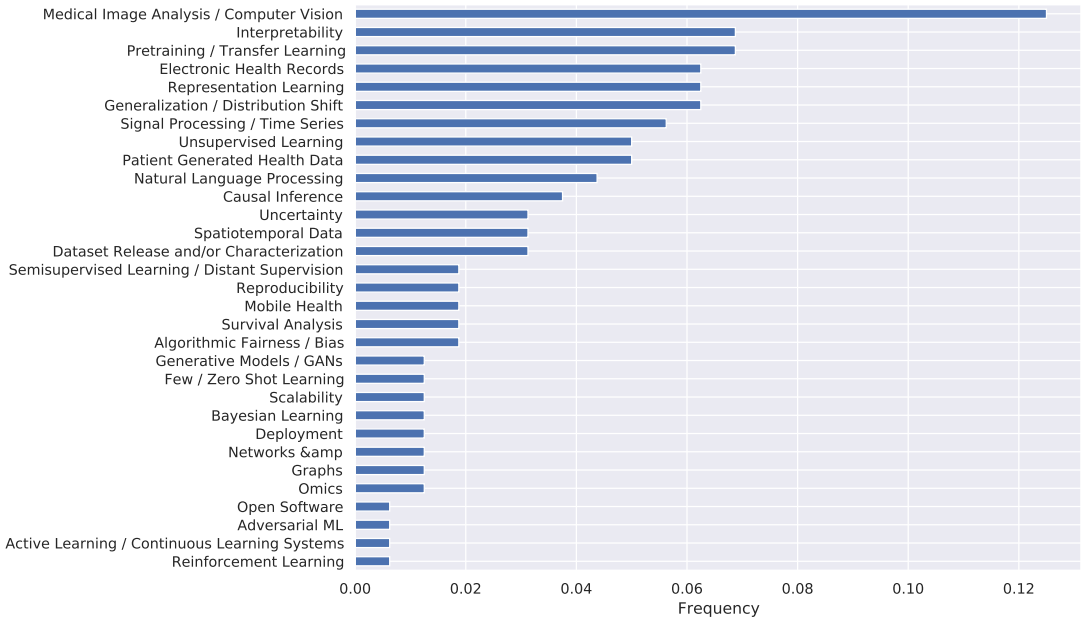

Figure 1: Frequency of author-reported topics for accepted papers.

3.2. Data and Methods

The authors were also asked to submit information about the statistical tests used for their analyses. Out of the accepted papers, 50.94% reported that variances/confidence intervals around point estimates were reported, 20.75% reported that statistical hypothesis testing or power analyses were performed in any model comparisons in addition to variance estimates, 15.09% outlined that only point estimates were computed and compared, and 13.20% reported that statistical analyses were not applicable for their paper. In addition, authors were asked whether the findings reported in the papers were validated across multiple healthcare sites. Among the accepted papers, 11.32% reported that results are demonstrated across multiple institutions by repeating independent training and evaluation on separate datasets and 9.43% answered that multi-site transfer-learning was assessed by training a model on one institution and evaluating on a second institution (with or without fine-tuning). Further details are provided in Table 1.

This year, authors were also asked to submit information about whether they compare model performance with baseline and/or SOTA models. Among the accepted papers, 43.39% reported that they compared results against previously published SOTA. Full statistics are reported in Table 2.

Authors of 73.58% of the accepted papers reported that they will make the source code of their work publicly available. In addition, among the accepted papers, 58.49% reported performing interpretability analysis.

3.3. Topic Modelling

We performed topic modeling over the free-text content of the accepted papers using Latent Dirichlet allocation (Blei et al., 2003). Both full proceedings papers and the extended abstracts were considered for this analysis. The marginal topic distribution is shown in Figure 2.

Table 1: Distribution of author responses on multi-site validation.

Option	%
No. Only one site's dataset was used.	41.51
Not Applicable	28.30
Yes, our results are demonstrated across multiple institutions by repeating independent training and evaluation on separate datasets.	11.32
Other	9.43
Yes, multi-site transfer-learning is assessed by training a model on one institution and evaluating on a second institution (with or without any second institution fine-tuning).	9.43

Table 2: Distribution of author responses on the comparison with baseline and/or SOTA models.

Option	%
Yes, we compared against previously published SOTA	43.39
Yes, we compared against handmade baselines (no prior published SOTA exists)	30.19
Not Applicable	13.21
No, we did not compare against baseline models	5.66
Yes, we compared against handmade baselines (though prior, published SOTA does exist)	5.66
Other	1.89

Compared to the topics from ML4H 2020 (Sarkar et al., 2020), new topics this year were "Robustness and Fairness", "ML in clinical trial design", and "ML in treatment and surgery". According to the model's classifications, there has been an increase in papers related to causal inference, robustness, and fairness which potentially points towards the growing interest in understanding the broader implications of the use of in machine learning for healthcare.

4. The ML4H Community

ML4H promotes the expansion of machine learning based health research across people from diverse backgrounds. Hence, we included a set of questions to better understand and improve the representation of diversity in this community and its accessibility to all members of the ML4H community. The symposium introduced submission, reviewer, and career mentorship programs, and research roundtables to encourage participants from underrepresented backgrounds to apply.

4.1. Experience

It was observed that 85.82% of all submissions and 79.24% of accepted papers originated from first authors who did not have a paper accepted at ML4H previously. In addition, the submitting authors were asked to state the approximate number of conferences/workshops which they have previously submitted their work to. 24.4% of authors submitted their first paper; 41.73%, 9.44%, and 13.38% have submitted 1-5, 5-10, and 10+ papers respectively, and the remainder chose not to report this information.

4.2. Gender

We requested the submitting author to optionally report their gender identity. We provided an inclusive set of options taken from the Iowa State Information and Data Collection Involving Gender and Sexuality such as agender, genderqueer, gender fluid, man, non-binary, questioning or unsure, transgender, trans man, trans woman, woman, and other. Moreover, we included options to allow the submitting author to provide a free-text input and also skip answering this question. 56.25% identified as man, 33.33% as woman, and the remaining 10.41% either preferred not to answer, was unknown, or non-binary. The submitting author was also asked to report whether any author in their paper identified with a gender that is underrepresented in machine learning research and 27.55% answered yes, 38.58% answered no, and 33.85% preferred not to answer or reported that it was unknown.

4.3. Race and Ethnicity

To better understand the racial and ethnic diversity of submissions, authors were also requested to optionally report their race and ethnicity. First, the submitting author was asked whether they belong to one of the following groups: Hispanic, Latino, or Spanish origin, such as Mexican, Puerto Rican or Cuban. Out of the submitting authors, 3.93% said yes, 74.01% said no, and 22.05% preferred not to answer. Second, regardless of the previous question, the submitting author was asked to report their race. A set of options were provided from Pew Research along with a field to add free-text entries. The distribution is shown in Figure 3 which shows that 40.62% of submitting authors identified as white, 37.5% as Asian or Asian American, 3.12% as Black or African, 2.08% as mixed, 2.08% as Middle Eastern, and 14.58% preferred not to answer. The submitting author was also asked to report whether any author in their paper identified with a racial group that is underrepresented

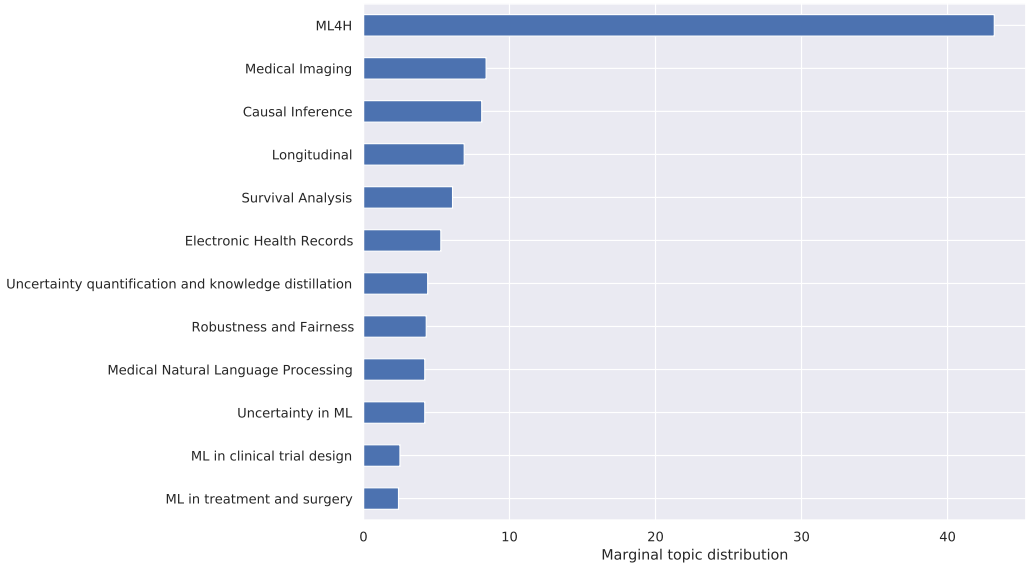

Figure 2: LDA topic distribution of accepted works in ML4H 2021.

in machine learning research and 14.17% answered yes, 49.60% answered no, and 36.62% preferred not to answer or reported that it was unknown.

4.4. Language

The submitting author community of ML4H represented a diverse language background. 51.96% of the submitting authors reported that they are non-native English speakers, 26.77% reported no, and 21.5% preferred not to answer.

4.5. Clinician Involvement

ML4H continues to seek high clinician participation. 66.13% of submitted papers involved at least a clinician either as primary author, secondary co-author, or consultant. The comparison with previous two years is shown in Table 3. While the overall clinician involvement has increased by 6.34% and 7.14% compared to ML4H 2020 and 2019 respectively, papers having a clinician as a primary author has reduced by 1.37% and 4.67% compared to ML4H 2020 and 2019 respectively.

5. Conclusion

The first Machine Learning for Health symposium (ML4H 2021) focused on highlighting the role of machine learning in health. We further reflected on how to improve access to our own community, through mentorship programs for submissions, reviewers, and career. The symposium featured a varied program of invited and spotlight talks, panels, and poster sessions. Despite the switch to a virtual format, interest in the intersection of machine learning and healthcare remains strong, and the symposium aimed to highlight and support this field and community.

6. Acknowledgements

6.1. Advisory Committee

Andrew Beam (Harvard Medicine, Faculty)
Brett K. Beaulieu-Jones (Harvard Medicine, Research Fellow)

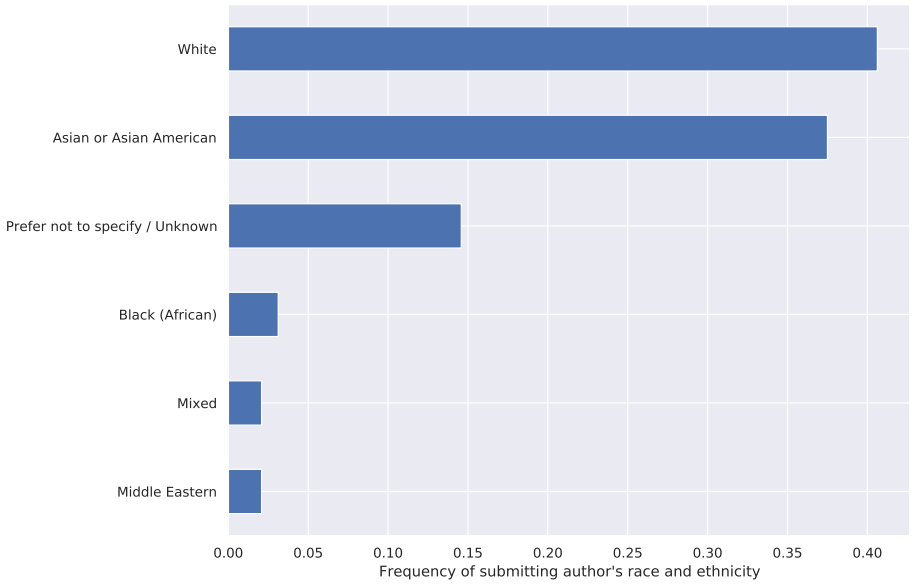

Figure 3: Frequency of reported submitting author races.

Table 3: Author responses on clinician involvement.

Clinician Involvement	2021	2020	2019
None	33.86%	40.2%	41.0%
Consultant/Acknowledged	20.47%	21.2%	19.1%
Secondary co-author	41.73%	33.3%	31.4%
Primary author	3.93%	5.3%	8.6%

Danielle Belgrave (Microsoft Research, Principal Research Manager)
Marzyeh Ghassemi (MIT, Assistant Professor)
Stephanie Hyland (Microsoft Research, Senior Researcher)

6.2. Meta reviewers

The organizers thank all meta-reviewers for their expert guidance in reviewing submissions to ML4H:

Neil Tenenholtz, Masoud Rouhizadeh, Sarah Tan, Craig Glastonbury, Liangqiong Qu, Edward Choi, Prithwish Chakraborty, Yuan Liu, Jessica Schrouff, George H. Chen, Sandhya Prabhakaran, Mehdi Fatemi, Negar Rostamzadeh, Yonatan Mintz, Mohammad Taha Bahadori, Shaun Canavan, Alexander Kotov, Di Jin, Emma Rocheteau, Brett Beaulieu-Jones, William Boag, Yanan Sui, Collin Stultz, Stefan Bonn, Ramasamy Savitha, Ayah Zirikly.

6.3. List of round table chairs

Senior roundtable chairs: Adarsh Subbaswamy, Wojciech Samek, Federico Cabitza, Lukas Ruff, Naomi Lee, Mark Sendak, Vidur Mahajan, Sergio Uribe, Jesse Ehrenfeld, Marzyeh Ghassemi, Polina Golland, David Buckeridge, Rumi Chunara, Judy Gichoya, Rohit Malpani, Andreas Reis, Jeff Lockhart, Roxana Daneshjou, Thomas J. Fuchs and Regina Geierhofer, Shada Al-Salamah

Junior roundtable chairs: Kaushik Manjunatha, Md. Golam Rasul, Utkarshani Jaimini, Huiqi Yvonne Lu, Kaivalya Deshpande, Sneha Jha, Rishab Khincha, Hang Yuan, Serifat Folorunso, Favour Nerrise, Jerry Fadugba, Elora Schörverth, Girmaw Abebe Tadesse, Rutwik Shah and Payal Chandak

6.4. List of reviewers

The proceedings and discussions would not have been possible without the generous reviewers who provided meaningful feedback and evaluation of all proceedings papers and extended abstracts. We are grateful to all reviewers:

Aahlad Puli, Aakanksha Rana, Aakash Kaku, Abdullah-Al-Zubaer Imran, Adam Yala, Ahmed Alaa, Alan D. Kaplan, Albee Ling, Alex Fedorov, Alexander Kotov, Ali Mottaghi, Ali Septiandri, Alison Callahan, Amy Lu, Anastasiia Kabeshova, Andrew Beam, Angeline Yasodhara, Anil Goyal, Aniruddh Raghu, Anirudh Joshi, Antonio Parziale, Aparna Balagopalan, Ari Ercole, Arjumand Masood, Arkabandhu Chowdhury, Arvind Pillai, Ashwin Vaswani, Asif Jan, Avik Ghose, Aviva Prins, Avnish Kumar, Ayah Zirikly, Azin Shamshirgaran, Benedek Rozemberczki, Boxiang Liu, Brett Beaulieu-Jones, Brian Cozzi, Byung-Hak Kim, Carson Drake, Chantal Shaib, Charles Delahunt, Chen Zhang, Cheryl Wong Sze Yin, Christina Ji, Christopher Bridge, Collin Stultz, Craig Glastonbury, Cristina Soguero-Ruiz, Dani Kiyasseh, David Belanger, David Zimmerer, Dawit Abtew, Deepak Sharma, Deepankar Nankani, Di Jin, Diana Mincu, Diwakar Mahajan, Doron Yaya-Stupp, Edward Choi, Edward De Brouwer, Elena Sizikova, Elham Dolatabadi, Elliot Schumacher, Emma Rocheteau, Eun-Kyeong Kim, Evan Cofer, Fabian Falck, Fabian Laumer, Farah Shamout, Fethi Jarray, Frank Rudzicz, Frank Soboczenski, Gagana B*, Gauri Deshpande, George H. Chen, Ghada Zamzmi, Giovanni Cina*, Girmaw Abebe Tadesse, Griffin Adams, Gulshan Sharma, Gurunath Reddy Madhumani, Hamed Javidi, Han Yuan, Hang Yuan, Hannes Stärk, Haoran Zhang, Harri Lähdesmäki, Hazrat Ali, Huiqi Yvonne

Lu, Hyunkwang Lee, Ioana Baldini, Irene Chen, Ivana Maric, Ivaxi Sheth, Jacob Deasy, Jan Witowski*, Jason Poulos, Jennifer Williams*, Jessica De Freitas, Jessica Schrouff, Jingshu Liu, Jocelyn Dunstan, John Zech, Jonas Kemp, Jonathan Deaton, Jongoh Jeong, Joongheon Kim, Jose Posada, Joseph Enguehard, Justin Lovelace, Juyong Kim, Ka-Ho Chow, Kajal Puri, Karin Stacke, Katharina Hoebel, Katikapalli Subramanyam Kalyan, Kaushik Sarveswaran, Kazuma Kobayashi, Ke Bai, Kexin Huang, Khawla Seddiki, Kuheli Sai, Kundan Krishna, Laura-Jayne Gardiner, Laya Rafiee Sevyeri, Layne Price, Liangqiong Qu, Liyue Shen, Lorenzo A. Rossi, Lovedeep Gondara, Mahmut Yurt, Malvern Madondo*, Manan Dey, Manish Gawali, Marta Avalos, Martin G Seneviratne, Masoud Rouhizadeh, Matt Barnes, Matthew Lungren, Matthew McDermott, Matthias Hüser, Matthias Kirchler*, Matthias Norden, Mayuree Binjolkar, Md Abdullah Al Hafiz Khan, Md Osman Gani, Md Taufeeq Uddin, Meghasyam Tummalacherla, Mehdi Fatemi, Mehreen Ali, Michael Dusenberry, Michael Oberst, Michal Borsky, Michelle Li, Min Lee, Mohammad Eslami, Mohammad Rasool Izadi, Mohammad Taha Bahadori, Mohammadreza Zandehshahvar, Monica Agrawal, Monica Munnangi, Mostafa 'Neo' Mohsenvand, Mozhgan Saeidi, Nandita Damaraju, Nazanin Makkinejad, Negar Rostamzadeh, Neil Tenenholtz, Nelson Johansen, Nick Phillips, Nicola Sahar, Niharika D'Souza, Nimshi Venkat Meripo, Nishamathi Kumaraswamy, Nishan Srishankar, Numan Celik, Olga Perepelkina, Oliver Rausch, Pablo Moreno-Muñoz, Pankaj Pandey, Partha Pratim Saha, Payal Chandak, Pedro Antonio Martin-Cervantes, Peng Liao, Petia Radeva, Pragyansmita Nayak, Praveen Gurunath Bharathi, Praveer Singh, Prithwish Chakraborty, Priyadip Ray, Qiao Jin, Qiyuan Hu, Rachid Riad, Rachneet Kaur, Rafael Sousa, Rafid Mahmood, Ramasamy Savitha, Rémy Dubois, Rishab Khincha, Rohan Mukherjee, Rohit Vashisht, Rutwik Shah, Ryan-Rhys Griffiths, Sadid Hasan, Saeed Alahmari, Sahra Ghalebikesabi, Sai Prabhakar Pandi Selvaraj, Sakyajit Bhattacharya, Salma EL Anigri, Sam Nguyen, Samaneh Kouchaki, Sami Mian, Sandhya Prabhakaran, Sandhya Tripathi*, Sangameshwar Patil, Sarah Brueningk, Sarah Tan, Sarvesh Nikumbh, Scott Fleming, Sebastian Caldas, Serina Y Chang, Seung Eun Yi, Sewale Demeke, Shalmali Joshi, Shanya Sharma, Shaun Canavan, Shreyasi Pathak, Simeon Spasov, So Yeon Min, Soumen Ghosh, Stefan Bonn, Stefan Groha*, Stefan Hegselmann, Stephanie Hyland, Stylianos Serghiou*, Subba Reddy Oota, Subhrajit Roy, Sudeshna Das, Sudhir Rachapalle, Suhani Vora, Sulaiman Somani, Sumit Mukherjee, Tanaya Babtiwale, Taylor Killian, Tejas Mathai, Thomas Y. Chen, Tianfan Fu, Tiffany Callahan, Tong Wu, Tony Y. Sun, Tristan Naumann, Tzu-Ting Wei, Urjoshi Sinha, Utkarshani Jaimini, Vaishnavi Subramanian, Victor Rodriguez, Vincent Jeanselme, Vincent Valton, Vinyas Harish, Vishwa Parekh, Vishwali Mhasawade, Viswesh Krishna, Vivek Natarajan, Vrutang Shah, Wangzhi Dai, Weicheng Zhu, Wenbin Zhang, Wenqi Wei, William Boag, William Carson, Wouter van Amsterdam, Xavier Sumba, Xiao Gu, Xiao Li, Xiaomeng Dong, Xiaoyi Zhang*, Xinyu Li, Yan Gao, Yan Zhuang, Yanan Sui, Yang Yang, Yannet Interian, Yi-han Sheu, Yiyang Yu, Yogatheesan Varatharajah, Yogesh Kumar, Yonatan Mintz, Young Ji Lee, Young Joon (Fred) Kwon, Yuan Liu, Yuchuan Wang, Yuwei Sun, Yuyin Zhou, Yvonne Lui, Zhaonan Sun, Zhengyuan Liu, Zhiwei Zhang, Zijun Zhang.

*denotes the top 10 reviewers who were recognized as excellent by the meta-reviewers. The criteria involved counting the number of

good reviews they have, normalize it against total number of papers they reviewed, then rank them by count of good reviews, breaking the tie with the normalized score.

References

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3: 993–1022, 2003.

Adrian V Dalca, Matthew BA McDermott, Emily Alsentzer, Samuel G Finlayson, Michael Oberst, Fabian Falck, Corey Chivers, Andrew Beam, Tristan Naumann, and Brett Beaulieu-Jones. Machine learning for health (ml4h) 2019: What makes machine learning in medicine different? In *Machine Learning for Health Workshop*, pages 1–9, 2020.

Fabian Falck, Yuyin Zhou, Emma Rocheteau, Liyue Shen, Luis Oala, Girmaw Abebe, Subhrajit Roy, Stephen Pfohl, Emily Alsentzer, and Matthew B. A. McDermott. A collection of the accepted abstracts for the machine learning for health (ml4h) symposium 2021, 2021.

Suproteem K. Sarkar, Subhrajit Roy, Emily Alsentzer, Matthew B. A. McDermott, Fabian Falck, Ioana Bica, Griffin Adams, Stephen Pfohl, and Stephanie L. Hyland. Machine learning for health (ml4h) 2020: Advancing healthcare for all. In Emily Alsentzer, Matthew B. A. McDermott, Fabian Falck, Suproteem K. Sarkar, Subhrajit Roy, and Stephanie L. Hyland, editors, *Proceedings of the Machine Learning for Health NeurIPS Workshop*, volume 136 of *Proceedings of Machine Learning Research*, pages 1–11. PMLR, 11 Dec 2020. URL https://proceedings.mlr.press/v136/sarkar20a.html.