Biomarkers AI and Bioinformatics Workshop Planning
Meeting #5

May 15, 2024

Agenda

1. Review task list

2. Advertising Slide for the EDRN Steering Committee Meeting

3. Session / Theme Chairs

4. Speakers

5. Posters

6. Session Organization

Proposed Tasks

1. ~~Goals and purpose~~
2. ~~High level sessions~~
3. ~~Workshop description drafted~~
4. ~~Identify additional program committee members~~
5. ~~Identify dates for the workshop which includes ensuring Caltech facilities are open~~
6. ~~Setup website on EDRN portal~~
7. ~~Distribute invitation and call for posters~~
8. Program committee member assignments (still open)
   - Session / Theme Chairs
   - Hack-a-thon Chairs
   - Invited speakers
9. Venue – meals and support (we will discuss this next time)
10. Review abstracts (May)
11. Skeleton structure (late May)
12. Open registration
13. Build schedule and continue to organize session and speakers

Sessions/Themes Assignments

Session/Themes	Organizer(s)	Moderator(s)	Speakers
In Silico and Real World Biomarker Discovery and AI	Zhen, Matt		
Biomarker Computation and Methodology Considerations for applying AI	Steve, Zhen		
Considerations in Data Preparation, Sharing, and Analysis	Eugene, Jen		
Emerging Capabilities and Methodologies in AI	Ashish, Matt		
Academic-Industry Partnerships in AI and Bioinformatics	Eugene, Chad		

Invited Speakers

- Prof. Changhuei Yang, Caltech, scanning systems optimized to collect the right kind of data for AI processing

- Prof. Bin Yu, UC Berkeley, data science, statistics and machine learning methods

- Prof. Faisal Mahood, Harvard University, AI for computational pathology and imaging

Posters

44 submissions with a decent distribution

We would like to add the program committee to help us in ensuring we should accept posters

    If you have a concern, please let us know

    For those accepted, we would like to see if any can be included as talks

Caltech is working to ensure we've to a space and poster boards for all the posters

For each session

Choose a moderator or set of moderators

- Structure session…suggestion but we are open:
  - Introductory framing of session by moderator - what are the objectives
  - Oral Talks
  - Moderated panel with questions

Sessions/Themes

In Silico and Real World Biomarker Discovery and AI

Cancer biomarker discovery — use cases and applications of AI

Biomarker data (knowledge) Integration: imaging, radiomics, high dimensional data (proteomics, genomics, epigenetics, etc.)

Showcasing datasets for AI/ML — real world data, synthetic data, and simulated data

Biomarker Computation and Methodology Considerations for applying AI

Uncertainty, statistical rigor, bias considerations in AI/ML

Statistical methods vs. AI: optimal approaches to enable data analysis

Trust in ML methodologies for cancer biomarkers: explainability, reproducibility, and interpretability of results

Considerations in Data Preparation, Sharing, and Analysis

Creating AI ready datasets: data preparation, harmonization, and standardization

Making data usable: data capture, sharing, federation, and scalable computation for cancer biomarkers

Using Open Science to link AI algorithms and datasets to publications

Ethical considerations with application of AI/ML

Emerging Capabilities and Methodologies in AI

Novel and emerging methods for image analysis

LLMs/Foundation Models, Generative AI, Federation/Federated Learning

Academic-Industry Partnerships in AI and Bioinformatics

Technology transfer

Structuring mutually beneficial relationships

Backup

Hackathon related timeline

Finalize hackathon planning team and setup call (April 19) - need a couple of volunteers

Finalize datasets (Tue, April 30) - need one more dataset

Problem definitions for hackathon (Tue, 21 May)

Techniques/libraries identified (Tue, 11 Jun)

Initial notebooks created (Tue, 2 Jul)

Datasets and problems and libraries and notebooks loaded (Tue, 23 Jul)

Meeting starts (Tue, 13 Aug)

Work continues (Aug 15++)

During the workshop

Day 1: Overview and problem description (poster?) - discussions around open problems / opportunities

Day 2: Informal discussions encouraged; rooms available for meetings

Hackathon Day 3: Short presentations: problem; data; focused exercises with AI ready datasets and configured notebooks

EDRN Datasets

Can use:

- Automated quantitative breast density measures (Moffitt; Dr. J Heine)
- Breast Cancer Biomarker Analysis (Moffitt; Dr. J Heine)

We are exploring a few other datasets but open to suggestions

Venue/Meals

Plan for morning and afternoon refreshments

Potential late afternoon/evening poster session and social

Lunch options

- Allow participants to pre-order for lunch brought in

- Lunch "on your own"