— AI/ML SECTOR OF THE EICUG SOFTWARE WORKING GROUP —

https://eic.ai/
https://indico.bnl.gov/e/AI4EIC

C. Fanelli, T. Horn

WILLIAM & MARY
CHARTERED 1693

powered by aws

What is AI4EIC?

http://www.eicug.org/content/wg.html

EIC has the unique opportunity to start incorporating AI from the very beginning and to systematically leverage on it during all phases of the project.

AI will be an integral part of the EIC software and we will take advantage of intelligent decisions in all aspects of data processing from detector readout and control to analysis.

To work in this direction, the AI4EIC working group (a sector of the SWG within the EICUG) has been established in Dec 2021. The AI WG (aka AI4EIC) serves as an entry point to AI applications and will organize workshops, tutorials, and Kaggle-like challenges.

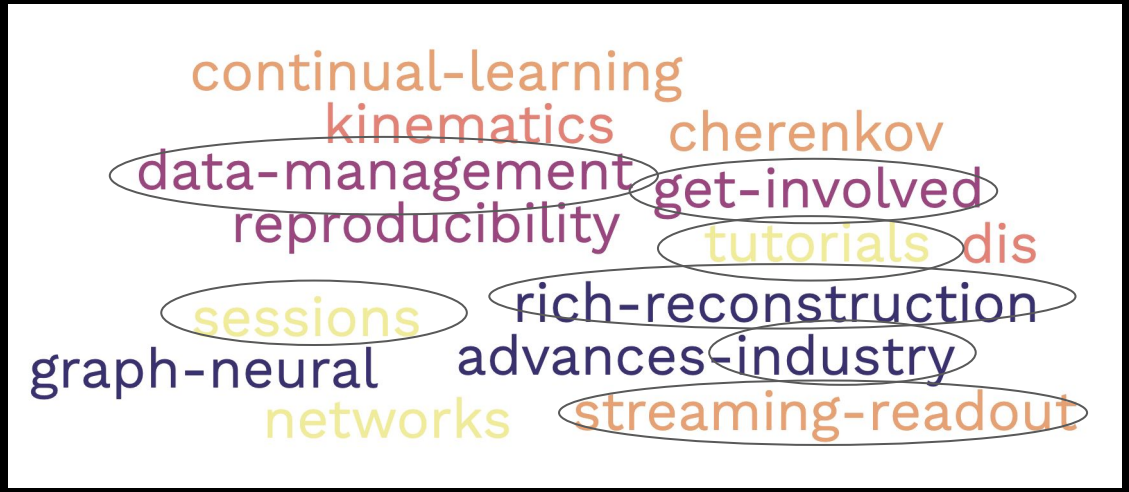
Educational activities are aimed at disseminating AI in the EIC community.

https://eic.ai

Survey (early 2022)

Hackathon: https://indico.bnl.gov/event/16586/page/435-hackathon
Tutorials: https://indico.bnl.gov/event/16586/page/426-tutorials

- A detailed survey was sent few months ago now https://forms.gle/6LADKTGaX7DeTVE46
- We want to learn more about our community, and we asked for feedback on what the needs and interests are, and what potential opportunities
- Feedback and key-words:



We organized our monthly meetings and this event (workshop and hackathon) taking into account this feedback

We have tutorial sessions every day of the workshop

For more details on the survey, see https://indico.bnl.gov/event/15636/

AI4EIC Events

The AIWG serves as an entry point to AI applications and will organize workshops, tutorials, and Kaggle-like challenges.

10/14/2022	**AI4EIC Hackathon** First AI4EIC hackathon, hosted by College of William and Mary, Williamsburg, VA October 14, 2022 https://doi.org/10.5281/zenodo.7197023 Le... + Event Details	In presence + virtual https://indico.bnl.gov/e/AI4EIC
10/10-13/2022	**AI4EIC Workshop** 2nd workshop, hosted by College of William and Mary, Williamsburg, VA October 10-13, 2022 For registration: https://indico.bnl.gov/e/AI4EIC	In presence + virtual https://indico.bnl.gov/e/AI4EIC
8/24/2022	**AI WG 5th meeting** 5th meeting of the EICUG AI WG topic-oriented: streaming readout + continual learning	**9am - 11am ET** Virtual https://indico.bnl.gov/event/16605/
7/20/2022	**AI WG 4th meeting** 4th meeting of the EICUG AI WG topic-oriented: detector design (re-scheduled; slides uploaded)	**9am - 11am ET** Virtual https://indico.bnl.gov/event/16328/
6/22/2022	**AI WG 3rd meeting** 3rd meeting of the EICUG AI WG topic-oriented: uncertainty quantification	**9am - 11am ET** Virtual https://indico.bnl.gov/event/16073

- Topic-oriented meetings: streaming readout, continual learning, detector design, uncertainty quantification
- Two well attended workshops (200++), that resulted in publications
- Hackathon

5/4/2022	**AI WG 2nd meeting** 2nd meeting of the EICUG AI WG Survey: https://forms.gle/6LADKTGqX7DeTVE46	**11am - 12pm ET** Virtual https://indico.bnl.gov/event/15636/
3/30/2022	**AI WG kickoff meeting** Kickoff meeting of the EICUG AI WG	**11am - 12pm ET** Virtual https://indico.bnl.gov/event/14923/
9/7-10/2021	**AI4EIC Workshop** 1st workshop , hosted by CFNS/BNL September 7-10, 2021	Virtual

https://eic.ai/events

AI4EIC 2nd Workshop: Oct 10-14 2022, William & Mary

- We had a total of 220 registered participants (also last year, >200!)
  - Very good attendance in person!
- Workshop: 6 sessions (15 conveners, 40+ speakers)
  - Design
  - Theory/Exp connections (morning + afternoon sessions)
  - Recon & PID
  - Infrastructure (+ Panel Discussion)
  - Streaming

- Educational: 4 Tutorials
  - MOBO, OmniFold, MLFlow, GNN

- Community: Hackathon (>90 interested people, 30+ participants from North, South America, Asia, Europe)

~300 users in ai4eic slack
sign-up here

AI4EIC Workshop (2022) Goals

- Goal of this second workshop was to cover all active and potential areas of applications of AI/ML for the EIC.
- Extend the scope of first workshop (focused on experimental applications only) to identify cross-cutting aspects and bring together different communities (accelerator, detector, theory, CS/DS)
- Identify impact and opportunities of AI in EIC
  - Write workshop report paper — comments from participants will be integrated (paper draft in progress; aim to complete by mid December)
  - The discussion from this workshop contributed to the LRP WP

AI4EIC Detailed Agenda

Convener list

Tutorial on Multi-objective Optimization (Meta)

Tutorial on Unfolding (LBNL, NERSC)

Mon, October 10, morning: introduction and overview

Mon, October 10, afternoon: Design

Tue, October 11, morning: theory/experiment connections

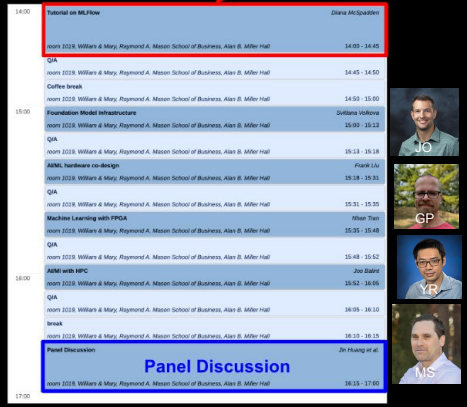
Tue, October 11, afternoon: theory/experiment connections

DOE and NSF Perspectives

Tutorial on MLFlow (JLab)

Wed, October 12, morning: Reconstruction & PID

Wed, October 12, afternoon: Infrastructure and Frontiers



Thu, October 13, morning: Streaming Readout

Tutorial on GNN (BNL)

Thu, October 13, afternoon: Summary and Future

AI4EIC Talks

Talks: https://indico.bnl.gov/event/16586/contributions/speakers

Design

- M. Balandat (Meta) Multi-objective Optimization Tutorial
- K. Suresh (Regina) Adaptive Experimentation in EIC
- B. Nachman (LBNL) AI-driven detector design
- E. Fol (CERN) ML Application for beam optics control in the LHC
- T. Satogata AI/ML overview for accelerator design activities

The/Exp (morn.)

- S. Liuti (UVA) ML for QCD analysis - 3D imaging
- T. Menzo (U. Cincinnati) Modeling Hadronization Using ML and the Lund String Model
- S. Andzrej (Jagiellonian U.) Modeling Hadronization Using ML and the Cluster Model
- A. Hiller Blin (U. Regensburg) A(I)DAPT
- N. Sato (JLab) Femtoscale Imaging of Nuclei using ML and Exascale Platforms
- B. Nachman (LBNL) Differentiable Simulations

Data centric analysis, UQ, modeling, analysis/preservation, event level inference

The/Exp (aftern.)

- D. Shih (Rutgers) Fast Detector Simulations with ML
- W. Phelps (CNU/JLab) ML in Spectroscopy and Partial Wave Analysis
- C. Pecar (Duke) Reconstructing DIS and SIDIS properties
- A. Butter (LPNHE CNRS) Ideas for ML based unfolding
- F. Torrales Acosta (LBNL) and V. Mikuni (NERSC) Unfolding Tutorial

Fast sim, spectroscopy, event-level reco, unfolding

Reco/PID Particle-level, event tagging, data-driven

- D. Whiteson (UC Irvine) Interpretable Networks for Identifying Leptons
- R. Kunnawalkam Elayavalli (Vanderbilt U.) Tagging heavy flavor jets @ RHIC
- W. Phelps (CNU/JLab) Muon identification with Deep Learning at EIC
- C. Allaire (IJC-Lab) Machine Learning in ACTS
- C. Peng (ANL) ML PID with measured shower profiles from calorimetry
- M. McEneaney (Duke) Λ event tagging at CLAS12
- N. Branson (Messiah U.), ML for calorimetry
- J. Giroux (Regina), Data-driven learning : Flux + Mutability

Infrastructure

- D. McSpadden (JLab/DS) MLFlow tutorial
- S. Volkova (PNNL) Foundation Model Infrastructure
- F. Liu (ORNL) AI/ML hardware co-design
- N. Tran (FNAL) Machine Learning with FPGA
- B. Joo (ORNL) AI/ML with HPC
- J. Huang (BNL), T. Miceli (FNAL), M. Williams (MIT), Panel Discussion

SRO

- M. Diefenthaler (JLab) INDRA-ASTRA
- S. Furletov (JLab) FastML for FPGA
- R. Ammendola (Tor Vergata, Rome) AI for streaming readout: an architectural perspective
- J. Huang (BNL) AI-based data reduction for streaming DAQ
- M. Bondi' (INFN/Catania) SRO for next generation electron scattering experiments
- C. Dean (MIT) ML for Heavy Flavor Identification
- T. Britton (JLab) AI for Experimental Controls

Live Document

https://docs.google.com/document/d/13BWnq_ywTYs__2zA0beeDob8pcuiiBiVqXNk33Kk8LY/edit

Live Document:

Artificial Intelligence for the Electron Ion Collider

Timetable

This is the live meeting notes (Q&A) document for the second workshop dedicated to Artificial Intelligence for the Electron Ion Collider, which will take place at William & Mary from **October 10 to October 14, 2022**.

This has been used during the workshop to collect questions and replies

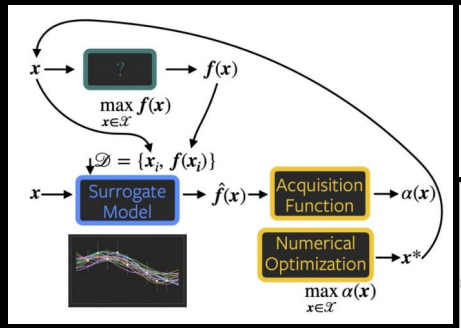
Conveners monitored discussion/questions in the live document

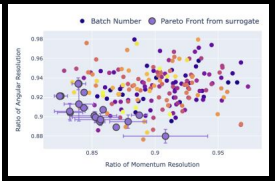
Total of 26 pages

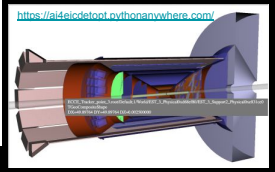
AI4EIC Tutorials

https://eic.ai/community

M. Balandat,
Meta AI



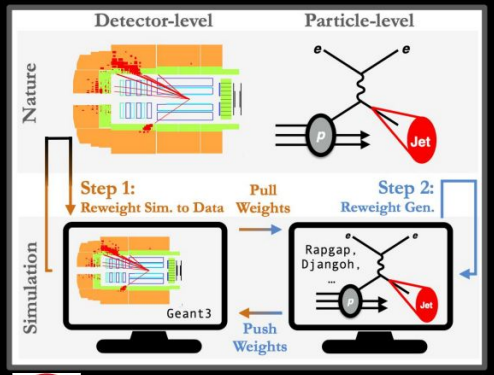




Multi-objective Optimization

F. Torales Acosta (LBNL)

V. Mikuni (NERSC)



OmniFold

Unfolding

K. Rajput
JLab/DS

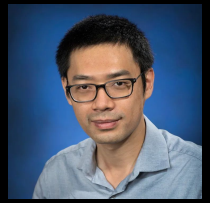
flow

Tracking
Record and query experiments: code, data, config, results

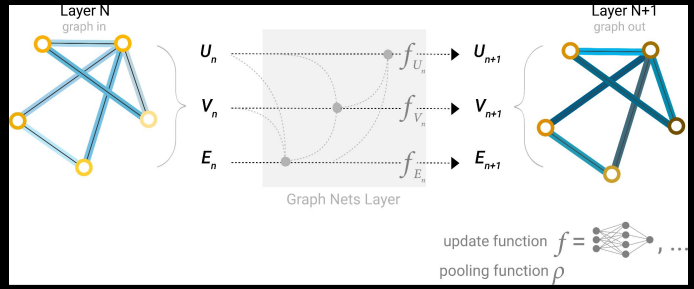
Projects
Packaging format for reproducible runs on any platform

Models
General format for sending models to diverse deploy tools

MLflow — ML lifecycle



Y. (Ray) Ren (BNL)



Graph Neural Network

AI4EIC Hackathon (Oct 14)

hackathon supported by AWS —
4 GPUs / instance on cloud computing;
1 instance / team;
10 teams total —
prize supported by W&M

powered by aws

10 teams, 30+ people, both in person and virtual; participation from America, Europe, Asia

- Documentation (problem description and dataset): https://doi.org/10.5281/zenodo.7197023
- Solutions accepted only above a certain threshold for the score
  - Hackathon winning team declared by 5pm ET of 10/14
    - Team JINR! (Alexey Aparin, Artem Korobitsin Grigorii Tolkachev, Vladimir Papoyan)
  - Solutions accepted for an additional week:
    - Best solution overall by Team Jets (Pierre Chatagnon Ibrahim Chahrour Dmitry Kalinkin Dulitha Jayakodige)

https://ai4eichackathon.pythonanywhere.com/leaderboard

AI4EIC Hackathon

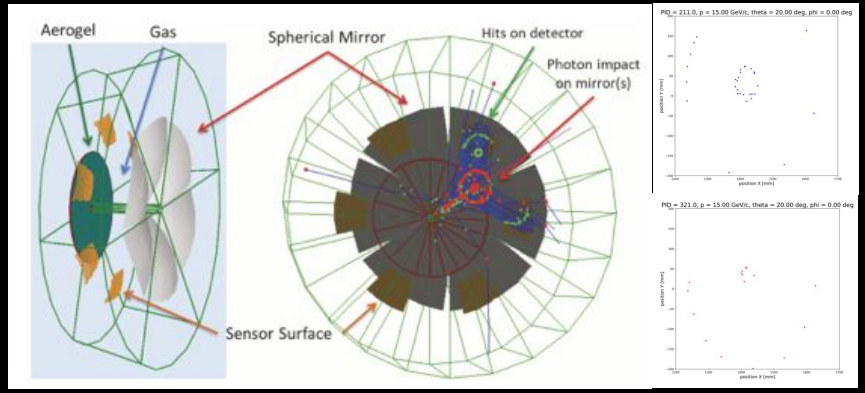
Congrats Team JINR!!!!!!!! (submission on 10-14-2022)

Hackathon Leaderboard

RANK	TEAM	SCORE	QUESTIONS ATTEMPED
1	Jets	295.502	Q 1, Q 3, Q 2
2	JINR	294.508	Q 1, Q 3, Q 2
3	JB and EC	262.313	Q 1, Q 3, Q 2

https://eic.ai/hackathons

Hackathon: Problems



Organizers: Cris Fanelli (William & Mary/JLab), Diana McSpadden (JLab/Data Science), Kishan Rajput (JLab/Data Science)

Advisory and problem definition: Evaristo Cisbani (INFN), Wouter Deconinck (U. Manitoba)

Computing resources: Eric Walter (William & Mary, IT)

Data generation, Documentation, Validation: James Giroux (U. Regina), Karthik Suresh (U. Regina)

Technical Assistance: Eric Walter (William & Mary, IT), James Giroux (U. Regina), Karthik Suresh (U. Regina)

CF, DMcS,KR, EC,WD,EW,KS,JG

Problem Number	Threshold Accuracy
Problem 1	94%
Problem 2	86%
Problem 3	80%

π, K datasets

Training Events	1.5 Million Events	With Magnetic Field (~ 1.5T)
Momentum	15 GeV/c	at Interaction Point (0, 0, 0)
Theta θ	20°	at Interaction Point (0, 0, 0)
Phi ϕ	0°	at Interaction Point (0, 0, 0)

Problem 1

Training Events	3 Million Events	With Magnetic Field (~ 1.5T)
Momentum	$15 - 20 GeV/c$	at Interaction Point (0, 0, 0)
Theta θ	$15 - 16°$	at Interaction Point (0, 0, 0)
Phi ϕ	$0 - 5°$	at Interaction Point (0, 0, 0)

Problem 2,3*

*Problem 3: addition of noisy hits

Solutions:

- JINR: CatBoost, https://catboost.ai/
- Jets: 2D CNN

The best solutions were all Machine Learning/Deep Learning-based, they were quite original, and they outperformed solutions based on classical approaches (followed by some teams). While this is only a first step towards deeply learning the identification of particles reconstructed with the dual-RICH, these exploratory studies clearly indicates the potential of ML/DL approaches for reconstruction and PID.

https://eic.ai/hackathons

AI4EIC Community

Based on the discussions during the workshop it is clear that AI4EIC fills an important need in the community by providing a forum for the community to come together and to learn from each other.

A few comments from workshop participants and the Panel session:

"AI4EIC very important to have in this growing community and to bring together people from different areas in the community"
"Need to keep moving things forward and AI4EIC provides a way to learn from each other and to broaden the community"
"Very fruitful to work with an interdisciplinary team…"
"Cannot be successful in the scientific communities alone. Used to go to many small workshops, which helps build connections."

A few overarching points that were discussed in more than one session

- ❑ Need to include expert/domain knowledge in the AI/ML approach and perhaps more importantly make clear that it is being included
- ❑ Need for encouraging participation of underrepresented groups in AI4EIC
- ❑ Need for the community work together in a coordinated way to be most efficient in reaching the overarching goals

AI4EIC Near Term

- Tools that have been important for AI4EIC and have received much community support

  - AI4EIC Topical Meetings via zoom approximately every month — they will be announced via multiple communication platforms such as webpage (https://eic.ai/events), mailing list (eicug-software-ai@eicug.org), and slack (https://ai4eic.slack.com/). Please join the mailing list to receive the announcements and forward to those you think should be on it, e.g., colleagues in theory or experiment
  - AI4EIC annual workshops — really brings together the community
    - Please send us an email if you are interested in hosting the AI4EIC workshop in 2023
    - Only requirement is that your institution can support an hybrid format
  - AI4EIC hackathons, kaggle-like challenges, etc

AI4EIC Longer Term

- Discussion of the possible realization of a venue for people to come together and share ideas, knowledge and resources

> “Several institutions could team up to provide alternative disciplines.”

- In addition to that, AI/ML activities are often fragmented, and despite this sparsity, similar efforts are often duplicated! Furthermore, AI/ML at scale and in the production environment, as well as dedicated heterogeneous resources, are limited. Need coordination and resources.

- Funding mechanism for AI/ML-related activities often complicated... Funding agencies presented several opportunities — need to discuss readiness with EIC PM, program managers, etc…

- AI4EIC looks timely given the data science expansion and emphasis.
  E.g., during the workshop, discussion on how this can increase involvement of universities, potential large educational component, etc.

Outlook

We plan to have the third annual workshop on Fall 2023

As discussed in the first AI4EIC workshop and hackathon locations will keep being rotated

Interested people can send an email with a proposal to support@eic.ai

More info on https://eic.ai/how-to-join