

4th ICFA Beam Dynamics Mini-Workshop on Machine Learning Applications for Particle Accelerators

Tuesday, March 5, 2024 - Friday, March 8, 2024

Lahan Select Gyeongju, South Korea

Book of Abstracts

Contents

Transformer Tutorial	1
Model up-keep with continual learning	1
Learning to Do or Learning While Doing: Reinforcement Learning and Bayesian Opti- mization for Online Continuous Tuning	2
Applying Reinforcement Learning to Particle Accelerators: An Introduction	2
Collaborations in ML: A Study in Scarlet	3
Physics Informed and Bayesian Machine Learning for Maximization of Beam Polarization at RHIC	4
Machine Learning to improve Accelerator Operations at SNS	5
Textual Analysis of ICALEPCS and IPAC Conference Proceedings: Revealing Research Trends, Topics, and Collaborations for Future Insights and Advanced Search	5
From Physics Study to Operations: Progress for Production Deployment of ML-Driven Beam Loss Optimization	6
Machine Learning Tools for Heavy-Ion Linac Operations	6
Controlling fusion plasmas with deep reinforcement learning	7
Bayesian Optimization with Neural Network Prior Mean Models	8
Multi-task Bayesian optimization of laser-plasma accelerators	8
A safe Bayesian optimization algorithm for tuning the optical synchronization system at European XFEL	9
From Simulation to Reality: Real-Time Control of Superconducting Linear Particle Accel- erator using a Trend-Based Soft Actor-Critic Algorithm	10
Machine learning to enhance XFEL operation at LCLS-II	10
Real-time Reinforcement Learning on FPGA with Online Training for Autonomous Accel- erators	11
Leveraging Vendor Tools for AI Acceleration	12
An Open Source MLOps Pipeline for Particle Accelerators	12

Lean MLOps stack for development and deployment of Machine Learning models into an EPICS Control system	13
Building a FAIR-Compliant Platform for AI-Ready Data in Particle Accelerators	13
Automation CERN –Progress with automating CERN’s accelerator fleet	14
Interpretable Machine Learning at the European XFEL	15
Real-time Integration of Machine Learning for Beam Size Control at the Advanced Light Source	15
A Front-End Framework with Embedded ML Tools for Automating Neutron Scattering Experiments	16
Data-Based Condition Monitoring and Disturbance Classification in Actively Controlled Laser Oscillators	17
Predicting preventable (slow) trips at NSLS-II	18
Beam Condition Forecasting with Non-destructive Measurements at FACET-II	18
Enhanced quench detection at the EuXFEL through a machine learning-powered approach	19
The L-CAPE Project at FNAL	19
Fast 6-Dimensional Phase Space Reconstructions using Generative Beam Distribution Models and Differentiable Beam Dynamics	20
High-dimensional characterization of coherent synchrotron radiation effects using generative-model-based phase space reconstruction methods	20
Detailed 4D phase space reconstruction of flat and magnetized beams using differentiable simulations and neural-networks	21
Interactive Visualization and Automated Analysis of Neutron Scattering Experiments	22
Computer vision for laboratory assistance tools	23
Towards Unlocking Insights from Logbooks using AI at DESY and BESSY	23
Flash Talks for Poster/Demos 4th ICFA MaLAPA	24
Integration of Multi-Objective Genetic Algorithm and neural networks in linac optimization.	24
Simulation methods of 3D coupled storage ring based on SLIM formalism	25
Online optimisations of lifetime and injection efficiency in the ESRF EBS storage ring	25
Machine Learning Based Response Matrix Correction	26
Orbit correction by machine learning in TPS storage ring	26
Comparing gradient-based and non-gradient modelling and optimisation methods for investigating synchrotron dynamics	27

Online image-based beam-dump anomaly detection	27
Multiobjective Optimization of Cyclotron Cavity Model using Neural Network	28
RL-Based Control Strategies for HIPI Accelerator	28
Reinforcement Learning Based Radiation Optimization at a Linear Accelerator	29
Dynamic vacuum and temperature predictions for informed anomaly detection at the CERN-SPS	30
Trust Region Bayesian Optimization for Online Accelerator Control	30
Anomaly Detection for Diode Failures	31
Multi objective Bayesian optimization of ECR ion source at the Linear IFMIF Prototype Accelerator.	31
Optimization of a Longitudinal Bunch Merge Gymnastic with Reinforcement Learning	32
Distance Preserving Machine Learning for Uncertainty Aware Accelerator Capacitance Predictions	32
Improving Surrogate Model Performance for Sparse Outputs in the Spatial Domain	33
Optimization design of photocathode injector assisted by deep Gaussian process	34
Development of beam transport system optimization method using VAE and Bayesian optimization	34
Experience with ML-driven applications at PETRA III	35
Rapid Tuning of Synchrotron Surrogate Model at the Recycler Ring	36
Analysis and Improvement of Generalisability of Anomaly Detection Methods	36
The Reinforcement Learning for Autonomous Accelerators International Collaboration	37
Simultaneous corrections of nonlinear errors in the LHC triplets using machine learning	38
Application of Machine Learning to Accelerator Operations at SACLA/SPring-8.	38
End-to-end Simulations and ML infrastructure for Light Sources	39
Improving Accelerator Surrogate Models with a Knowledge of Physics	39
Reshaping SRF Cavity Resonance Management with Smart Techniques	40
Reinforcement Learning for Intensity Tuning at Large FEL Facilities	41
Research on Recognition of Quench and Flux Jump Based on Machine Learning	41
Towards Natural Language-driven Autonomous Particle Accelerator Tuning	42
Bayesian Optimal Experimental Design for AGS Booster Magnet Misalignment Estimation	43
Improving Electronic Logbook Searches Using Natural Language Processing	43

PACuna: Automated Fine-Tuning of Language Models for Particle Accelerators	44
AI/ML Coupling & Surrogates in BLAST Accelerator Modeling Codes	44
Experience with ML-based Model Calibration Methods for Accelerator Physics Simulations	45
ML methods for noise reduction in industrial LLRF systems	46
Leveraging Differential Algebraic Methods for Enhanced Beam Dynamics Simulation with Machine Learning	47
Cheetah –A High-speed Differentiable Beam Dynamics Simulation for Machine Learning Applications	47
AI-assisted design of final cooling system for a muon collider	48
4th ICFA MaLAPA Summary	48
Closing	49

Tutorials / 93**Transformer Tutorial****Author:** Anton Lu¹¹ *Technical University of Vienna (AT)***Corresponding Author:** anton.lu@cern.ch

The Transformer is a deep learning architecture introduced in 2017, that has since then taken over the natural language processing field and has recently gained public popularity thanks to large language models like ChatGPT. The self-attention mechanism introduced with the Transformer allows it to learn complex patterns and relationships in data without explicitly using recurrent mechanisms like classic RNN-style architectures. While the Transformer was developed for sequence-to-sequence language modeling like translation tasks, the usefulness for time series prediction has been less explored in the machine learning community. Particularly, the lack of beginner-friendly tutorials and guides for using transformers with uni- and multivariate continuous input and outputs are not easily found online, as opposed to for natural language tasks. Therefore, this tutorial aims to introduce the Transformer architecture and how to use standard deep-learning library Transformer building blocks to construct a simple time series prediction model and explain the inputs and outputs of the transformer model. As an appendix, we will give a quick outlook of current state-of-the-art time series prediction architectures based on the basic Transformer as well as alternative modern time series forecasting methods.

Primary Keyword:

transformers

Secondary Keyword:**Tertiary Keyword:****Tutorials / 79****Model up-keep with continual learning****Author:** Kishansingh Rajput¹**Co-authors:** Malachi Schram ¹; Ricardo Vilalta ²¹ *Jefferson Lab*² *University of Houston***Corresponding Authors:** schram@jlab.org, kishan@jlab.org, r.vilalta.us@ieee.org

Particle accelerators are dynamic machines and pose a major challenge for scientists at the intersection of nuclear physics and machine learning (ML) with evolving operational conditions and data drift. Traditional ML models trained on historical data can fail to provide good predictions on future data. They fall short in adapting to dynamic distributions. This tutorial introduces the particle accelerator community to continual learning techniques to address this challenge. The tutorial covers fundamentals of concept/data drift, drift detection, continual learning, online learning for model-upkeep, transfer learning along with potential practical use cases.

Primary Keyword:**Secondary Keyword:****Tertiary Keyword:**

Tutorials / 34

Learning to Do or Learning While Doing: Reinforcement Learning and Bayesian Optimisation for Online Continuous Tuning

Authors: Chenran Xu¹; Jan Kaiser²

Co-authors: Andrea Santamaria Garcia ³; Annika Eichler ⁴; Erik Bründermann ⁵; Florian Burkart ⁴; Frank Mayet ⁴; Hannes Dinter ⁴; Holger Schlarb ⁴; Oliver Stein ⁴; Thomas Vinatier ⁴; Willi Kuropka ⁴

¹ *Karlsruhe Institut für Technologie (KIT)*

² *Deutsches Elektronen-Synchrotron DESY*

³ *Karlsruhe Institute of Technology*

⁴ *DESY*

⁵ *Karlsruhe Institute of Technology (KIT)*

Corresponding Authors: erik.bruendermann@kit.edu, annika.eichler@desy.de, jan.kaiser@desy.de, andrea.santamaria@kit.edu, chenran.xu@kit.edu

Online tuning of particle accelerators is a complex optimisation problem that continues to require manual intervention by experienced human operators. Autonomous tuning is a rapidly expanding field of research, where learning-based methods like Bayesian optimisation (BO) hold great promise in improving plant performance and reducing tuning times. At the same time, Reinforcement Learning (RL) is a capable method of learning intelligent controllers, while recent work shows that RL can also be used to train domain-specialised optimisers in so-called Reinforcement Learning-trained Optimisation (RLO). In parallel efforts, both algorithms have found successful adoption in particle accelerator tuning. Here we present a comparative case study, analysing the behaviours of both algorithms and outlining their strengths and weaknesses. The results of our study help provide criteria for choosing a suitable learning-based tuning algorithm for a given task and will accelerate research and adoption of these methods with particle accelerators and other complex real-world facilities, ultimately improving their availability and pushing their operational limits, thereby enabling scientific and engineering advancements.

Primary Keyword:

ML-based optimization

Secondary Keyword:

reinforcement learning

Tertiary Keyword:

bayesian optimization

Tutorials / 75

Applying Reinforcement Learning to Particle Accelerators: An Introduction

Authors: Andrea Santamaria Garcia¹; Chenran Xu²; Jan Kaiser³

¹ *Karlsruhe Institute of Technology*

² *Karlsruhe Institut für Technologie (KIT)*

³ *Deutsches Elektronen-Synchrotron DESY*

Corresponding Authors: jan.kaiser@desy.de, andrea.santamaria@kit.edu, chenran.xu@kit.edu

Reinforcement learning is a form of machine learning in which intelligent agents learn to solve complex problems by gaining experience. In current research, agents trained with reinforcement learning perform better than their human counterparts on problems that have historically been difficult for machines to solve. Particle accelerators are among the most advanced high-tech machines in the world. Modern scientific experiments place the highest demands on beam quality, making particle accelerator control extremely complex. Reinforcement learning is a promising avenue of research that has the potential to improve existing accelerator control solutions and enable new ones that have previously been impossible with conventional methods. The barrier of entry into reinforcement learning, however, is high and slows its adoption in the accelerator field. In this tutorial, we apply reinforcement learning to the task of tuning transverse beam parameters in a real-world accelerator beam line and focus in particular on solving the issues that arise in the context of particle accelerators, such as the high cost of samples, a large sim2real gap and the high non-linearity of the control and optimisation tasks under investigation.

Primary Keyword:

reinforcement learning

Secondary Keyword:

AI-based controls

Tertiary Keyword:

ML-based optimization

Field Summaries / 51

Collaborations in ML: A Study in Scarlet

Author: Joshua Einstein-Curtis¹

Co-authors: Jonathan Edelen ¹; Rebecca Cotter ¹; Wil Jacques ²

¹ *RadiaSoft LLC*

² *Foresight Science and Technology*

Corresponding Authors: jedelen@radiasoft.net, joshec@radiasoft.net, wil.jacques.foresightst@gmail.com, cotter@radiasoft.net

Robust, stable collaborations are a challenge to create in highly technical fields, particularly in situations where funding streams are unreliable. Difficulties can be found in technical, organizational, and legal spheres; these can include: language differences, differing data formats, and legacy work. Creating a collaboration for machine learning in accelerators is particularly challenging due to the variety and unclear scope of member interests. This includes the volume of data and data sharing, model development and sharing, and workflow sharing. There is also a significant need to develop a trained, expert workforce. Presented here are the preliminary results of research about ML usage in US accelerator facilities, as well as challenges facing any large-scale software collaborations in the government sphere. This includes examinations of collaboration models already in the space, such as used in control systems and experimental data sharing, and funding and organizational agreements that have seen success.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Accelerator R&D and Production, under Award Number DE-SC0024543.

Primary Keyword:

MLOps

Secondary Keyword:

Tertiary Keyword:

Field Summaries / 84

Physics Informed and Bayesian Machine Learning for Maximization of Beam Polarization at RHIC

Authors: Auralee Edelen¹; Bohong Huang²; Christopher Hall³; David Sagan⁴; Georg Hoffstaetter^{None}; Jonathan Edelen³; Kevin Brown^{None}; Malachi Schram⁵; Natalie Isenberg⁶; Nathan Cook³; Nathan M. Urban⁶; Vincent Schoefer^{None}; Weijian Lin⁴; Yinan Wang⁷; Yuan Gao^{None}

¹ SLAC, Stanford

² Stony Brook University

³ RadiaSoft LLC

⁴ Cornell University

⁵ Thomas Jefferson National Accelerator Facility

⁶ Brookhaven National Laboratory

⁷ Rensselaer Polytechnic Institute

Corresponding Authors: bohong.huang@stonybrook.edu, nurban@bnl.gov, dcs16@cornell.edu, wangy88@rpi.edu, ygao@bnl.gov, brownk@bnl.gov, schoefer@bnl.gov, georg.hoffstaetter@cornell.edu, nisenberg@bnl.gov, wl674@cornell.edu, ncCook@radiasoft.net, chall@radiasoft.net, jedelen@radiasoft.net, schram@jlab.org, edelen@slac.stanford.edu

The Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory (BNL) provides the world's only high-energy polarized proton beam. It is in the unique position to study from where nuclei obtain their spin, which is ultimately the property responsible for any kind of magnetization. Preserving polarization during particle acceleration is a difficult process that requires tuning many accelerator components. RHIC's successor, the Electron Ion Collider (EIC) will be one of the most complex scientific instruments ever built, with the capability of colliding polarized proton and electron beams. This increase in instrument complexity will require new, sophisticated tools to optimize accelerator performance thereby maximising the utility of polarized beam experiments.

A collaborative project between BNL, JLab, Cornell, RPI, SLAC, and RadiaSoft is tackling the challenging problem of polarization maximization for RHIC. The polarization benefits from three intermediate objectives: (1) preserve beam density, (2) synchronize accelerator components at depolarizing resonance crossings, and (3) minimize depolarizing resonance strengths. Operational parameters for these objectives can be measured and actuated, making them suitable candidates for optimization. We aim to build a fully Bayesian, physics-informed machine learning (ML) framework to optimize the system and maximize polarization performance. Because the EIC will use the same polarized pre-accelerator chain as RHIC, this methodology will improve future EIC polarization performance as well.

In this presentation, we will describe the project's objectives, initial results, and future plans. Major components of the project include improving physics-based models of the accelerator and learning accurate physics- and data-driven surrogate models by leveraging ML-based model calibration methods. We will use these models in conjunction with Bayesian optimization, reinforcement learning, and fast feed-forward ML-based corrections to achieve optimal polarization performance at RHIC.

Primary Keyword:

ML-based optimization

Secondary Keyword:

digital twins

Tertiary Keyword:

AI-based controls

Field Summaries / 78**Machine Learning to improve Accelerator Operations at SNS****Author:** Kishansingh Rajput¹**Co-authors:** Malachi Schram ¹; Willem Blokland ²; Alexander Zhukov ²; Charles Peters ²; Pradeep Ramuhalli ²; Ricardo Vilalta ³¹ *Jefferson Lab*² *Oak Ridge National Laboratory*³ *University of Houston***Corresponding Authors:** r.vilalta.us@ieee.org, zhukovap@ornl.gov, blokland@ornl.gov, peterscc@ornl.gov, ramuhallip@ornl.gov, kishan@jlab.org, schram@jlab.org

Particle accelerators are made up of thousands of components with many pieces of equipment running at their peak power. As a consequence, particle accelerators can fault and abort operations for numerous reasons. In order to avoid these faults, we apply uncertainty aware Machine Learning (ML) based anomaly prediction techniques. We predict any unusual behavior and perform preemptive actions in order to improve the total availability of particle accelerators. One of the challenges with application of ML on particle accelerators is the variability of the data due to changes in the system configurations. To address this, we employ conditional models. In addition, distance-based uncertainty awareness allows us to decide when a model need to be re-trained/tuned for continual learning with drift in the data. In this talk, we present an overview of various ML use-cases being explored at Spallation Neutron Source (SNS) accelerator to improve efficiency. Further, we present errant beam prognostics in detail including experimental setup, data collection, curation, model training, evaluation and deployment results. In addition, we present comparison between semi-supervised (conditional variational auto-encoder) and supervised learning (conditional Siamese model) methods for anomaly detection at SNS.

Primary Keyword:

failure prediction

Secondary Keyword:

uncertainty quantification for ML

Tertiary Keyword:

anomaly detection

Field Summaries / 57**Textual Analysis of ICALEPCS and IPAC Conference Proceedings: Revealing Research Trends, Topics, and Collaborations for Future Insights and Advanced Search****Authors:** Annika Eichler¹; Antonin Sulc²; Tim Wilksen²¹ *Deutsches Elektronen Synchrotron DESY*

² DESY MCS**Corresponding Authors:** antonin.sulc@desy.de, tim.wilksen@desy.de

In this paper, we show a textual analysis of past ICALEPCS and IPAC conference proceedings to gain insights into the research trends and topics discussed in the field. We use natural language processing techniques to extract meaningful information from the abstracts and papers of past conference proceedings. We extract topics to visualize and identify trends, analyze their evolution to identify emerging research directions, and highlight interesting publications based solely on their content with an analysis of their network. Additionally, we will provide an advanced search tool to better search the existing papers to prevent duplication and easier reference findings. Our analysis provides a comprehensive overview of the research landscape in the field and helps researchers and practitioners to better understand the state-of-the-art and identify areas for future research.

Primary Keyword:**Secondary Keyword:****Tertiary Keyword:****Optimization & Control / 80**

From Physics Study to Operations: Progress for Production Deployment of ML-Driven Beam Loss Optimization

Author: Carrie Elliott¹¹ ORNL**Corresponding Author:** elliottcl@ornl.gov

C. Elliott, W. Blokland, D. Brown, B. Cathey, B. Maldonado Puente, C. Peters, K. Rajput, J. Rye, M. Schram, S. Thomas, A. Zhukov

Spallation Neutron Source, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

*Thomas Jefferson National Laboratory, Newport News, VA, 23606, USA

The Spallation Neutron Source (SNS) at Oak Ridge National Laboratory (ORNL), a high-power H-

*ORNL is managed by UT-Battelle, LLC, under contract DE-AC05-00OR22725 for the U.S. Department of Energy.

Primary Keyword:

ML-based optimization

Secondary Keyword:**Tertiary Keyword:****Optimization & Control / 70**

Machine Learning Tools for Heavy-Ion Linac Operations

Authors: Brahim Mustapha¹; Jose Martinez-Marin¹¹ Argonne National Laboratory

Corresponding Author: brahim@anl.gov

At a heavy ion linac facility, such as ATLAS at Argonne National Laboratory, a new ion beam is tuned once or twice a week. The use of machine learning can be leveraged to streamline the tuning process, reducing the time needed to tune a given beam and allowing more beam time for the experimental program. After establishing automatic data collection and two-way communication with the control system, we have developed and deployed machine learning models to tune and control the machine. We have successfully trained online different Bayesian Optimization (BO)-based models for different sections of the linac, including the commissioning of a new beamline. We have demonstrated transfer learning from one ion beam to another allowing fast switching between different ion beams. We have also demonstrated transfer learning from a simulation-based model to an online machine model and using Neural Networks as the prior-mean for BO optimization. Following a failed attempt to deploy Reinforcement Learning (RL), we have finally succeeded in training a model online for one beam and deploy it for the tuning of other beams. More recently, these models are being generalized to other sections of the ATLAS linac and can, in principle, be adapted to control other ion linacs and accelerators with modern control systems.

- This work was supported by the U.S. Department of Energy, under Contract No. DE-AC02-06CH11357. This research used the ATLAS facility, which is a DOE Office of Nuclear Physics User Facility.

Primary Keyword:

ML-based optimization

Secondary Keyword:

bayesian optimization

Tertiary Keyword:

reinforcement learning

Keynote / 35

Controlling fusion plasmas with deep reinforcement learning

Author: Jaemin Seo¹

¹ *Chung-Ang University*

Corresponding Author: jseo@cau.ac.kr

The tokamak is one of the most promising concepts for confining fusion plasma. Controlling the tokamak actuators to stably maintain plasma in the desired state is an essential technology for sustainable energy production using nuclear fusion. Recently, technologies controlling fusion plasma in the tokamak using deep reinforcement learning (RL) have been emerging. In this presentation, we will present research results on optimizing the actuation trajectory, controlling the plasma state, and maintaining the plasma stability in tokamak devices using deep RL.

Primary Keyword:

AI-based controls

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

reinforcement learning

Optimization & Control / 16

Bayesian Optimization with Neural Network Prior Mean Models

Author: Tobias Boltz¹

Co-authors: Auralee Edelen¹; Brahim Mustapha²; Daniel Ratner¹; Jose Martinez-Marin²; Ryan Roussel¹

¹ *SLAC National Accelerator Laboratory*

² *Argonne National Laboratory*

Corresponding Authors: tboltz@slac.stanford.edu, rroussel@slac.stanford.edu, edelen@slac.stanford.edu

Bayesian optimization using Gaussian processes is a powerful tool to automate complex and time-consuming accelerator tuning tasks and has been demonstrated to outperform conventional methods at several facilities. In high-dimensional input spaces, however, even this sample efficient search may take a prohibitively large number of steps to reach convergence. In this contribution, we discuss the use of neural networks as a prior mean to inform the surrogate GP model and thereby speed-up convergence. We present collaborative results obtained in simulations and experiments at the Linac Coherent Light Source (LCLS) and the Argonne Tandem Linear Accelerator System (ATLAS). We show that high quality models can significantly improve optimization performance and discuss further measures to recover performance in cases where only models of limited accuracy are available.

Primary Keyword:

bayesian optimization

Secondary Keyword:

Tertiary Keyword:

Optimization & Control / 19

Multi-task Bayesian optimization of laser-plasma accelerators

Author: Remi Lehe¹

¹ *LBNL*

Corresponding Author: rlehe@lbl.gov

When designing a laser-plasma acceleration setup, it is common to explore the parameter space (plasma density, laser intensity, focal position, etc.) with Particle-In-Cell (PIC) simulations in order to find an optimal configuration that, for example, minimizes the energy spread or emittance of the accelerated beam. However, PIC simulations can be computationally expensive. Various reduced models (e.g., reduced-model simulation codes, ML surrogates) can approximate beam behavior at a much lower computational cost. Although such models do not capture the full physics, they could still suggest promising sets of parameters to be simulated with a full PIC code and thereby speed up the overall design optimization.

In this work, we automate such a workflow with a Bayesian multitask algorithm, where the different tasks correspond to simulations making different approximations. This algorithm learns from past

simulation results with different fidelities, and dynamically chooses the next parameters to be simulated. The libEnsemble library is used to orchestrate this workflow on a modern GPU-accelerated high-performance computing system.

Primary Keyword:

bayesian optimization

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

MLOps

Optimization & Control / 24

A safe Bayesian optimization algorithm for tuning the optical synchronization system at European XFEL

Author: Jannis Lübsen¹

Co-authors: Annika Eichler²; Maximilian Schütte²; Sebastian Schulz²

¹ TUHH

² DESY

Corresponding Authors: annika.eichler@desy.de, jannis.luebsen@tuhh.de, maximilian.schuette@desy.de

Over recent years, Bayesian optimization has become a widely adopted tool for fine-tuning and enhancing the operational performance of particle accelerators. While many Bayesian optimization (BO) algorithms focus on unconstrained optimization, constraints play an important role in accelerator operations. They ensure the safe functioning of the equipment and prevent damage to expensive components. Frequently, these constraints are not known and must be acquired through learning. In Sui et al., a safe Bayesian optimization method was introduced. This method actively learns safety constraints during the optimization process, ensuring a safe operating environment. To tackle high-dimensional optimization challenges, LineBo was introduced to decompose the overall optimization domain into smaller sub-domains.

In this research, we present a modification of the SafeOpt approach that separates exploration and exploitation. This proposed strategy accelerates the convergence rate in particle accelerator applications, particularly at the European XFEL, the world's largest linear particle accelerator. The European XFEL is renowned for its capacity to produce intense and ultra-short X-ray flashes for investigating ultra-fast, time-resolved chemical processes. To enhance the quality of observations, the laser-based synchronization system is fine-tuned using PI controllers, employing a safe Bayesian optimization technique.

Given the high cost of machine time at the European XFEL, it is imperative for the algorithm to identify optimal parameters as fast as possible. In our contribution, we introduce a safe Bayesian optimization algorithm that not only ensures safety but also significantly improves convergence speed and noise robustness. We illustrate its application and present comparative results through simulations for the optical synchronization system of the European XFEL, as well as an experimental demonstration on a laboratory synchronization system.

Primary Keyword:

bayesian optimization

Secondary Keyword:

active learning

Tertiary Keyword:

ML-based optimization

Optimization & Control / 2

From Simulation to Reality: Real-Time Control of Superconducting Linear Particle Accelerator using a Trend-Based Soft Actor-Critic Algorithm

Author: Xiaolong Chen¹

Co-author: Chunguang Su²

¹ *Institute of modern physics*

² *Institute of modern physics, Chinese Academy of Sciences*

Corresponding Authors: chenxiaolong@impcas.ac.cn, suchunguang@impcas.ac.cn

Superconducting linear accelerators play a vital role in advancing scientific discoveries by requiring frequent reconfiguration and tuning. Minimizing setup time is crucial to maximize experimental time. Recently, reinforcement learning (RL) algorithms have emerged as effective tools for solving complex control tasks across various domains. Nonetheless, deploying RL agents trained in simulated environments to real-world scenarios remains a challenge. To address this, we propose a novel paradigm for transferring RL agents from simulation to real accelerators. In this study, we introduce the trend-based soft actor-critic (TBSAC) method and showcase its effectiveness through two successful applications in real-world linear particle accelerators. TBSAC demonstrates strong robustness, allowing agents trained in simulated environments to be applied to real-world accelerators. We validated our method by performing a typical beam control tasks at the China Accelerator Facility for Superheavy Elements (CAFe II). In the orbit correction tasks at CAFe II, our approach reduces the tuning time conducted by human experts by tenfold, achieving a corrected orbit with RMS values of less than 1mm. These experiments clearly demonstrate the efficiency and effectiveness of our proposed approach while maintaining expert standards. Our method holds significant potential for future applications in accelerator commissioning fields.

Primary Keyword:

AI-based controls

Secondary Keyword:

reinforcement learning

Tertiary Keyword:

ML-based optimization

Optimization & Control / 81

Machine learning to enhance XFEL operation at LCLS-II

Author: Zihan Zhu¹

Co-authors: Auralee Edelen²; Ryan Roussel¹; Daniel Ratner³; Dylan Kennedy³; Eric Cropp³; Tobias Boltz¹

¹ *SLAC National Accelerator Laboratory*

² *SLAC, Stanford*

³ SLAC

Corresponding Authors: tboltz@slac.stanford.edu, zihanzhu@slac.stanford.edu, rroussel@slac.stanford.edu, edelen@slac.stanford.edu, cropp@slac.stanford.edu, dratner@slac.stanford.edu

The LCLS-II is a high repetition rate upgrade to the Linac Coherent Light Source (LCLS). LCLS-II will provide up to a million pulses per second to photon science users. The emittance and dark current are both critical parameters to optimize for ideal system performance. The initial commissioning of the LCLS-II injector was substantially aided by detailed online physics modeling linking high performance computing directly to physicists in the control room, along with the use of Bayesian optimization for fine tuning the emittance while balancing against dark current.

Here we summarize the role these tools played in the commissioning period and are playing in the current operational stage of the LCLS-II injector, which provides an example for how other accelerator facilities may benefit from combining online modeling and optimization infrastructure. We also describe current progress on creating a fully deployed digital twin of the LCLS-II injector based on a combination of ML modeling and physics modeling, using the LUME software suite and various ML-based characterization tools. Finally, we will describe current efforts and plans to leverage the online LCLS-II injector model in fast optimization and control schemes.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Optimization & Control / 23

Real-time Reinforcement Learning on FPGA with Online Training for Autonomous Accelerators

Author: Luca Scomparin¹

Co-authors: Andrea Santamaria Garcia ²; Andreas Kopmann ¹; Anke-Susanne Müller ¹; Chenran Xu ³; Edmund Blomley ¹; Erik Bründermann ⁴; Johannes L. Steinmann ¹; Jürgen Becker ¹; Michele Caselle ¹

¹ KIT

² Karlsruhe Institute of Technology

³ Karlsruhe Institut für Technologie (KIT)

⁴ Karlsruhe Institute of Technology (KIT)

Corresponding Authors: edmund.blomley@kit.edu, juergen.becker@kit.edu, erik.bruendermann@kit.edu, michele.caselle@kit.edu, andreas.kopmann@kit.edu, luca.scomparin@kit.edu, anke-susanne.mueller@kit.edu, johannes.steinmann@kit.edu, andrea.santamaria@kit.edu, chenran.xu@kit.edu

Reinforcement Learning (RL) is a promising approach for the autonomous AI-based control of particle accelerators. Real-time requirements for these algorithms can often not be satisfied with conventional hardware platforms.

In this contribution, the unique KINGFISHER platform being developed at KIT will be presented. Based on the novel AMD-Xilinx Versal platform, this system provides cutting-edge general microsecond-latency RL agents, specifically designed to perform online-training in a live environment.

The successful application of this system to dampen horizontal betatron oscillations at the KARlsruhe Research Accelerator (KARA) will be discussed. Additionally, preliminary results of the application of the system to the highly non-linear problem of controlling microbunching instabilities will be presented.

Primary Keyword:

reinforcement learning

Secondary Keyword:

AI-based controls

Tertiary Keyword:

Infrastructure / Deployment Workflows / 52

Leveraging Vendor Tools for AI Acceleration

Authors: Joshua Einstein-Curtis¹; Nathan Cook¹

Co-authors: Curtis Berger ²; Jeroen van Tilborg ²; Jonathan Edelen ¹; Rob Nagler ¹; Samuel Barber ²; Stephen Coleman ¹

¹ *RadiaSoft LLC*

² *Lawrence Berkeley National Lab*

Corresponding Authors: curtisberger@lbl.gov, coleman@radiasoft.net, ncook@radiasoft.net, jvantilborg@lbl.gov, sbarber@lbl.gov, joshec@radiasoft.net, jedelen@radiasoft.net, nagler@radiasoft.net

Several large vendors have been expanding their ML deployment tooling to allow for easy deployment of machine learning models on processing devices. AMD Xilinx has developed a toolkit for accelerating ML calculations on their FPGAs by utilizing either dedicated “AI Engine”(AIE) hardware or an openly available IP block known as “Deep Learning Processing Units”(DPUs). Vitis AI is actively maintained, with regular version releases. Google has created extensions to TensorFlow to compile models to programs that can run on dedicated accelerators, such as those provided by Coral.ai, while also allowing for other deployment methods. This is known as TensorFlow Lite, part of the TensorFlow ecosystem and TFX deployment pipelines. This toolchain is able to compile for ARM, Xilinx, and GPUs. Presented here is the use of one of these toolchains to develop a laser focal position controller at LBNL’s BELLA facility, including a discussion of future plans for easing control and deployment needs for such a system.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number(s) DE-SC0021680 and Prime Contract No. DE-AC02-05CH11231.

Primary Keyword:

AI-based controls

Secondary Keyword:

MLOps

Tertiary Keyword:

Infrastructure / Deployment Workflows / 4

An Open Source MLOps Pipeline for Particle Accelerators

Author: Gopika Bhardwaj¹

¹ *Fermilab*

Corresponding Author: gopikab@fnal.gov

In this talk we will describe the open-source MLOps (Machine Learning Operations) framework tools selected and designed to support ML (Machine Learning) algorithm development and deployment on Fermilab's main accelerator complex. MLOps is the standardization and streamlining of the ML development lifecycle to address the challenges and risks associated with large-scale machine learning projects such as: changing data dependencies, varying operations needs, reproducibility, and diverse teams working with differing tools and skills. We will show how this framework is being used for minimizing the injection and extraction losses of our Booster synchrotron.

Primary Keyword:

MLOps

Secondary Keyword:

Tertiary Keyword:

Infrastructure / Deployment Workflows / 62

Lean MLOps stack for development and deployment of Machine Learning models into an EPICS Control system

Author: Mateusz Leputa¹

¹ *UKRI-STFC-ISIS*

Corresponding Author: mateusz.leputa@stfc.ac.uk

The ISIS Neutron and Muon Source is undergoing several upgrades to the control hardware, software, data acquisition and archiving systems. Machine learning systems are also being integrated into the control system. This not only requires the models to be high-quality but also to be maintained and kept up to date, especially in performance-critical applications. Each model incurs additional code-based maintenance, which can affect a project's longevity. For these reasons, we implemented a workflow for training and deploying models that utilises off-the-shelf, industry-standard tools such as MLflow, TF-Serve and Torch-Serve. We discuss the use of these tools; the adoption of lean paradigms and DevOps practices help optimise the developer throughput, maximise model quality and simplify monitoring and retraining. These tools and practices minimise the developer time spent on non-ML tasks, make it easy to track and compare model changes and performance as well as help improve the overall visibility of the projects across ML teams. We demonstrate how the models are automatically deployed and integrated into the EPICS control system and served to the end-user via Phoebus controls, which decreases the turn-around time for user feedback and together with the centralised model storage helps deliver and iterate the models rapidly. We discuss challenges and lessons learned along the development process, as well as a direction for future developments.

Primary Keyword:

MLOps

Secondary Keyword:

AI-based controls

Tertiary Keyword:

Infrastructure / Deployment Workflows / 29

Building a FAIR-Compliant Platform for AI-Ready Data in Particle Accelerators

Authors: Xiaohan Lu^{None}; Yi Jiao¹

¹ *Institute of High Energy Physics*

Corresponding Authors: jiaoyi@ihep.ac.cn, luxh@ihep.ac.cn

In the rapidly evolving field of particle accelerator technology, Machine Learning (ML) shows great potential in optimizing accelerator performance and predictive maintenance. However, the success of these applications often depends on high-quality, real-time data sources. This paper introduces the plan and status of building an innovative machine learning data acquisition platform, specifically designed for continuously generating machine learning datasets applicable in the field of particle accelerators. The platform adheres to the FAIR principles, i.e., the Findability, Accessibility, Interoperability, and Reusability of the data. It also will ensure data integrity, consistency, and reliability during data collection, ultimately generating an AI-ready dataset. By utilizing advanced data collection networks, edge computing, and storage technologies, the platform achieves real-time data capture, preprocessing, and annotation. Importantly, the platform will include a variety of algorithms that intelligently select and store data points based on real-time operational states and the needs of machine learning models, and provides appropriate annotations for the data. This research will not only offer a powerful data support tool for the operation and maintenance of particle accelerators but also provide new perspectives and methods for the industrial sector on how to effectively acquire and manage data for machine learning applications.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Infrastructure / Deployment Workflows / 8

Automation CERN –Progress with automating CERN’s accelerator fleet

Author: Michael Schenk¹

Co-authors: Francesco Velotti¹; Verena Kain¹

¹ *CERN*

Corresponding Authors: francesco.maria.velotti@cern.ch, verena.kain@cern.ch, michael.schenk@cern.ch

Automation has become one of the key topics of preparing CERN’s accelerators for the future. It has been identified as essential for the existing complex to cope with upcoming challenges as well as for future projects such as the FCC, that will require a significant reduction of exploitation cost compared to today’s standards to get accepted. In recent years, the highest impact areas of automation have been identified and a roadmap for implementation has been prepared. It should allow full automation of most standard processes. This contribution will show recent progress with autonomous control building blocks and results in operation. Organizational aspects and new interesting algorithms will also be discussed. And finally the remaining questions and challenges will be mentioned.

Primary Keyword:

AI-based controls

Secondary Keyword:

MLOps

Tertiary Keyword:

reinforcement learning

Infrastructure / Deployment Workflows / 55

Interpretable Machine Learning at the European XFEL

Author: Danilo Enoque Ferreira de Lima¹

Co-authors: Amna Majida¹; Luca Gelisio¹

¹ *European X-ray Free Electron Laser*

Corresponding Authors: danilo.enoque.ferreira.de.lima@xfel.eu, luca.gelisio@xfel.eu, amna.majida@xfel.eu

The success of experiments at large scale photon sources is strongly connected with the quality of collected data and the information available to scientists during the beamtime. Similarly, streamlined and automated operation of the facility can minimize inefficiencies, thereby boosting the scientific outcome.

The key strategic goal of the Machine Learning program at the European XFEL is to empower scientists through strict information quality control and easily explainable metrics. Each application developed provides a reliability estimate, either through the usage of carefully estimated uncertainties or test statistics designed to verify the trustworthiness of the results. Additionally, the choice of algorithm is guided by a scientific methodology applicable to the situation at hand. The deployment of such methods takes advantage of Karabo, the control system at the European XFEL, which offers a unified point of entry for scientists to monitor and steer the instruments.

In this presentation, we introduce selected applications resulting from the program. We use Bayesian optimization to tune analysis parameters online for serial femtosecond crystallography, based on scientifically relevant metrics. A similar procedure exploits information theory for the automated multi-modular geometry optimization. Another activity exploits machine learning to automate and enhance X-ray spectral diagnostics. Furthermore, a self-supervised approach is used to automatically characterize collected data, so as to highlight interesting data samples. A final contribution includes the research of predictive maintenance methods to identify faults early and react fast, with the aim of preventing a cascade failure and maximizing beamtime efficiency.

The Machine Learning program at the EuXFEL aims to streamline and automate operation of the facility to minimize inefficiencies and boost scientific outcome, and it has been established as a key asset at the European XFEL. Its successful implementation can only be achieved through a strategy of open communication of methods and quality control.

Primary Keyword:

bayesian optimization

Secondary Keyword:

uncertainty quantification for ML

Tertiary Keyword:

failure prediction

Infrastructure / Deployment Workflows / 60

Real-time Integration of Machine Learning for Beam Size Control at the Advanced Light Source

Author: Thorsten Hellert¹

Co-authors: Andrea Pollastro ; Hiroshi Nishimura ; Marco Venturini ; Simon C. Leemann ; Tynan Ford

¹ *Lawrence Berkeley National Laboratory*

Corresponding Authors: scleemann@lbl.gov, mventurini@lbl.gov, apollastro@lbl.gov, h_nishimura@lbl.gov, tford@lbl.gov, thehellert@lbl.gov

The Advanced Light Source (ALS) storage ring employs various feedback and feedforward systems to stabilize the circulating electron beam thus ensuring delivery of steady synchrotron radiation to the users.

In particular, active correction is essential to compensate for the significant perturbations to the transverse beam size induced by user-controlled tuning of the insertion devices, which occurs continuously during normal operation. Past work at the ALS already offered a proof-of-principle demonstration that Machine Learning (ML) methods could be used successfully for this purpose.

Recent work has led to the development of a more robust ML-algorithm capable of continuous re-training and its routine deployment into day-to-day machine operation. In this contribution we focus on technical aspects of gathering the training data and model analysis based on archived data from 2 years of user operation as well as on the model implementation including the interface of an EPICS Input/Output Controller (IOC) into a Phoebus Panel, enabling operator-level supervision of the Beam Size Control (BSC) tool during regular user operation. This deployment ensures real-time integration of machine learning models into the ALS control system.

Primary Keyword:

MLOps

Secondary Keyword:

active learning

Tertiary Keyword:

ML-based optimization

Infrastructure / Deployment Workflows / 45

A Front-End Framework with Embedded ML Tools for Automating Neutron Scattering Experiments

Authors: Matthew Kilpatrick¹; Morgan Henderson¹

Co-author: Jonathan Edelen¹

¹ *RadiaSoft LLC*

Corresponding Authors: mhenderson@radiasoft.net, jedelen@radiasoft.net, kilpatrick@radiasoft.net

The rscontrols framework developed at RadiaSoft was created to simplify controls automation for neutron scattering experiments using machine learning (ML), beginning with sample alignment. Written in Python, rscontrols uses virtual representations of equipment and controls to enable seamless integration of hardware, EPICS protocols, and analytical tools including deep networks and other ML models. Embedded UNet image segmentation models have already been deployed through rscontrols in live tests at ORNL in which we have successfully demonstrated automated sample alignment. Significant effort has been dedicated to live uncertainty quantification for generating

reliability metrics and diagnostics, retaining the security of human feedback while reducing reliance on input for operations. In addition to image segmentation, a UNet encoder/decoder has been used along with a non-ML filter method for live neutron camera image denoising. Implementations of additional ML models are underway, including Bayesian optimization and reinforcement learning schemes to further automate controls and optimize the scientific value of scattering data. This will incorporate work from parallel efforts at RadiaSoft to perform detector-based fine-tuning using reciprocal space data and a 3D UNet.

Primary Keyword:

AI-based controls

Secondary Keyword:

uncertainty quantification for ML

Tertiary Keyword:

MLOps

Anomaly Detection / Failure Prediction / 9

Data-Based Condition Monitoring and Disturbance Classification in Actively Controlled Laser Oscillators

Author: Arne Grünhagen¹

Co-authors: Annika Eichler²; Marina Tropmann-Frick¹; Görschwin Fey³

¹ *Hamburg University of Applied Sciences*

² *Deutsches Elektronen Synchrotron DESY*

³ *Hamburg University of Technology*

Corresponding Author: arne.gruenhagen@desy.de

The successful operation of the laser-based synchronization system of the European X-Ray Free Electron Laser relies on the precise functioning of numerous dynamic components operating within closed loops with controllers. This study presents a comprehensive overview of the application of data-driven machine learning methods to detect and classify disturbances in these dynamic systems, leveraging the output signals from controllers. Four distinct feature extraction techniques are introduced, encompassing statistical analysis in both the time and frequency domains, characteristics of spectral peaks, and the use of autoencoder-generated latent space representations in the frequency domain. Remarkably, these methods do not necessitate system-specific knowledge and can be adapted for deployment in other dynamic systems. This research integrates feature extraction, fault detection, and fault classification into an automated and comprehensive condition monitoring framework. To achieve this, a systematic comparison is undertaken, evaluating the performance of 19 state-of-the-art fault detection algorithms and four classification algorithms. The objective is to identify the most suitable combination of feature extraction and fault detection or classification algorithms for effectively modeling the condition of an actively controlled phase-locked laser oscillator. Experimental evaluations show the effectiveness of clustering algorithms, highlighting their capacity to detect perturbed system conditions. Furthermore, our evaluation shows the support vector machine as the most suitable choice for classifying different types of disturbances in the laser-based synchronization system.

Primary Keyword:

anomaly detection

Secondary Keyword:

Tertiary Keyword:

Anomaly Detection / Failure Prediction / 91

Predicting preventable (slow) trips at NSLS-II

Author: Reid Smith¹

¹ *BNL*

Corresponding Author: smithr@bnl.gov

NSLS-II has been working with SLAC and Argonne on ML applications for improving accelerator reliability; specifically in predicting preventable (slow) trips & in using anomaly detection to identify most likely trip-causes to reduce recovery time. We are several years into the project, and already have positive results in the ‘trip prevention’ application.

Primary Keyword:

anomaly detection

Secondary Keyword:

Tertiary Keyword:

Anomaly Detection / Failure Prediction / 13

Beam Condition Forecasting with Non-destructive Measurements at FACET-II

Author: Matthew Kilpatrick¹

Co-authors: Brendan O’Shea ²; Paul Moeller ¹; Robbie Watt ²

¹ *RadiaSoft LLC*

² *SLAC*

Corresponding Authors: boshea@slac.stanford.edu, moeller@radiasoft.net, rwatt@slac.stanford.edu, kilpatrick@radiasoft.net

Beam diagnostic technology is one of the foundations of large particle accelerator facilities. A challenge with operating these systems is the measurement of beam dynamics. Many methods such as beam position monitors have an inherent destructive quality to the beam and produce perturbations after the measurement. The ability to measure the beam conditions with non-destructive edge radiation allows for us to have a more stable understanding and predictability of the beam condition. We are developing a machine learning workflow for the downstream prediction and future forecasting of the beam condition utilizing the non-destructive edge radiation measurements and novel graph neural networks in collaboration with FACET-II at SLAC. Our methods divide the problem into two different aspects. First, we are developing machine learning algorithms with the beam physics integrated within each layer of the network. Second, developing an online surrogate model of edge radiation using SRW will allow for automatic generation of new beam states due to the changing parameters of accelerator facilities over time. We plan to integrate and test our prediction system at the SLAC facility to perform beam condition prediction and verification at FACET-II.

Primary Keyword:

failure prediction

Secondary Keyword:

surrogate model architecture

Tertiary Keyword:

uncertainty quantification

Anomaly Detection / Failure Prediction / 66

Enhanced quench detection at the EuXFEL through a machine learning-powered approach

Author: Lynda Boukela¹

Co-authors: Annika Eichler¹; Julien Branlard¹; Nur Zulaiha Jomhari¹

¹ *German Electron Synchrotron (DESY)*

Corresponding Authors: annika.eichler@desy.de, nur.zulaiha.jomhari@desy.de, julien.branlard@desy.de, lynda.boukela@desy.de

Within the context of the European X-Ray Free-Electron Laser (EuXFEL), where 800 superconducting radio-frequency cavities (SRFCs) are employed to accelerate electron bunches to energies as high as 17.5 GeV, ensuring safe and optimal accelerator operation is crucial. In this work, we introduce a machine learning (ML)-enhanced approach for detecting anomalies, with a particular focus on identifying quenches, which can disrupt the superconductivity of the SRFCs, leading to operational interruptions. Our method consists of a two-stage analysis of the cavity dynamics. We first leverage analytical redundancy to process the data, and generate a residual for statistical testing and anomaly detection. Subsequently, we employ machine learning to distinguish quenching events from other anomalies. Different algorithms have been explored, and adapted in order to take into account the specificity of the data at hand. The evaluation, based on 2022 data, demonstrates the superior performance of our approach when compared to the currently deployed quench detection system.

Primary Keyword:

anomaly detection

Secondary Keyword:

Tertiary Keyword:

Anomaly Detection / Failure Prediction / 69

The L-CAPE Project at FNAL

Author: Brian Schupbach¹

¹ *FNAL*

Corresponding Author: schupbach@fnal.gov

The Linac Condition Anomaly Prediction Emergence Project (L-CAPE) at Fermilab National Accelerator Lab (FNAL) seeks to apply data-analytic methods to improve the information available to MCR Operators and to automate the task of labeling Linac outage types as they occur by recognizing patterns in real-time machine data. Predicting outages in a credible manner could provide useful information for managing the impact of the outage on the other accelerators in the complex thereby minimizing downtime and leading to potential energy savings. An overview of the methods and challenges of gathering machine data via the existing Accelerator Controls system for training, developing, and deploying an ML model will be discussed.

Primary Keyword:

failure prediction

Secondary Keyword:

timeseries forecasting

Tertiary Keyword:

MLOps

Analysis & Diagnostics / 22

Fast 6-Dimensional Phase Space Reconstructions using Generative Beam Distribution Models and Differentiable Beam Dynamics

Authors: Auralee Edelen¹; Juan Pablo Gonzalez-Aguilera²; Ryan Roussel¹

¹ *SLAC National Accelerator Laboratory*

² *University of Chicago*

Corresponding Authors: jpga@uchicago.edu, edelen@slac.stanford.edu, rroussel@slac.stanford.edu

Next-generation accelerator concepts, which hinge on the precise shaping of beam distributions, demand equally precise diagnostic methods capable of reconstructing beam distributions within 6-dimensional phase spaces. However, the characterization of intricate features within 6-dimensional beam distributions using conventional diagnostic techniques necessitates hundreds of measurements, using many hours of valuable beam time. Novel diagnostic techniques are needed to substantially reduce the number of measurements required to reconstruct detailed, high dimensional beam features as feedback in precision beam shaping applications. In this study, we present a novel approach to analyzing experimental measurements using generative machine learning models of 6-dimensional beam distributions and differentiable beam dynamics simulations. We demonstrate in simulation that using our analysis technique, conventional beam manipulations and diagnostics can be used to reconstruct detailed 6-dimensional phase space distributions using as few as 20 beam measurements with no prior training or data collection. These developments enable detailed, high dimensional phase space information as online feedback for precision control of beam distributions in advanced accelerator applications and can be used to improve our understanding of complex accelerator beam dynamics.

Primary Keyword:

transformers

Secondary Keyword:

differentiable models

Tertiary Keyword:

Analysis & Diagnostics / 28

High-dimensional characterization of coherent synchrotron radiation effects using generative-model-based phase space reconstruction methods

Authors: Juan Pablo Gonzalez-Aguilera¹; Ryan Roussel²; Auralee Edelen²; Young-Kee Kim¹

¹ *University of Chicago*² *SLAC National Accelerator Laboratory***Corresponding Authors:** jpga@uchicago.edu, rroussel@slac.stanford.edu, edelen@slac.stanford.edu, ykkim@hep.uchicago.edu

Coherent synchrotron radiation (CSR) is a detrimental effect in linear accelerators due to its contribution to projected emittance growth and microbunching. However, conventional measurement techniques are not precise enough to resolve the exact multi-dimensional effects of CSR, namely the different rotation of transverse phase space slices throughout the longitudinal coordinate of the bunch. In this work, we investigate the applicability of our generative-model-based six-dimensional phase space reconstruction method in the detailed characterization of CSR effects at the Argonne Wakefield Accelerator Facility. Additionally, we study the current resolution limitations of the phase space reconstruction method and perform an analysis of its accuracy and precision in simulated cases. Finally, we test the reconstruction algorithm with synthetic beams that approximate distributions affected by CSR.

Primary Keyword:

differentiable models

Secondary Keyword:**Tertiary Keyword:****Analysis & Diagnostics / 5**

Detailed 4D phase space reconstruction of flat and magnetized beams using differentiable simulations and neural-networks

Authors: Seongyeol Kim¹; John Power²; Philippe Piot³; Scott Doran²; Wanming Liu²; Chen Gongxiaohui²; Eric Wisniewski²; Charles Whiteford²; Auralee Edelen⁴; Ryan Roussel⁴; Juan Pablo Gonzalez Aguilera⁵; Inhyuk Nam⁶; Myunghoon Cho¹; Chi Hyun Shim¹; Haeryong Yang¹; Gyujin Kim¹; Kookjin Moon¹; Sunghoon Kwon¹; Chang-Kyu SUNG¹

¹ *Pohang Accelerator Laboratory*² *Argonne National Laboratory*³ *Northern Illinois University*⁴ *SLAC National Accelerator Laboratory*⁵ *University of Chicago*⁶ *Pohang Accelerator Laboratory (PAL)*

Corresponding Authors: highlong@postech.ac.kr, ilyoukim@postech.ac.kr, ihnam@postech.ac.kr, mh0309@postech.ac.kr, sch0914@postech.ac.kr, csung@postech.ac.kr, kookjine@postech.ac.kr, jp@anl.gov, cwhiteford@anl.gov, wmliu@anl.gov, ppiot@anl.gov, sdoran@anl.gov, rroussel@slac.stanford.edu, jpga@uchicago.edu, sykim12@postech.ac.kr, edelen@slac.stanford.edu, ewisniew@anl.gov, b288079@anl.gov, sh.kwon@postech.ac.kr

Phase space reconstruction method based on differentiable simulations and neural-networks [R. Roussel et al., Phys. Rev. Lett. 130, 145001, 2023] is robust diagnostics to map complete 4D (x, x', y, y') phase space. It was reported that this method provides not only the uncoupled phase space ($x-x'$, and $y-y'$), but also coupled phase space information ($x-y'$, and $y-x'$). Recently, in addition to round, uncorrelated beam, we experimentally demonstrated the phase space reconstruction of flat and magnetized beams at the Argonne Wakefield Accelerator facility. Here, flat beam indicates the large emittance ratio in horizontal and vertical planes (e.g., $\epsilon_{nx}/\epsilon_{ny} \gg 100$), and magnetized beam represents the transversely coupled beam due to canonical angular momentum from non-zero magnetic field at the cathode. In this study, we show that the phase space reconstruction method also provides the complete 4D phase space of those special beams such that we can capture i) the large emittance ratio of the flat beam and ii) magnetization value of the magnetized beam. In addition, we will discuss the uncertainty of the beam parameters obtained from conventional diagnostic method and phase space reconstruction.

Primary Keyword:

differentiable models

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

uncertainty quantification

Analysis & Diagnostics / 12**Interactive Visualization and Automated Analysis of Neutron Scattering Experiments****Author:** Matthew Kilpatrick¹**Co-authors:** Alexander Kuhn²; Andrei Savici³; Bogdan Vacaliuc³; Christina Hoffman³; David Bruhwiler¹; Dragos Tatulea²; Evan Carlin¹; Gurhar Khalsa¹; Jorg Mensmann²; Marc Nienhaus²; Matthew Tucker³; Peter Messmer²; Steffen Roemer²; Zachary Morgan³¹ *RadiaSoft LLC*² *NVIDIA*³ *Oak Ridge National Laboratory***Corresponding Authors:** mnienhaus@nvidia.com, sromer@nvidia.com, tuckermg@ornl.gov, dtatulea@nvidia.com, bruhwiler@radiasoft.net, pmessmer@nvidia.com, akuhn@nvidia.com, gurhar@radiasoft.net, saviciat@ornl.gov, choffmann@ornl.gov, jmensmann@nvidia.com, kilpatrick@radiasoft.net, vacaliucb@ornl.gov, morganzj@ornl.gov, evan@radiasoft.net

Neutron scattering experiments have undergone significant technological development through large area detectors with concurrent enhancements in neutron transport and electronic functionality. Data collected for neutron events include detector pixel location in 3D, time and associated metadata, such as, sample orientation, neutron wavelength, and environmental conditions. RadiaSoft and Oak Ridge National Laboratory personnel are considering single-crystal diffraction data from the TOPAZ instrument. We are leveraging a new method for rapid, interactive analysis of neutron data using NVIDIA's IndeX 3D volumetric visualization framework. We have implemented machine learning techniques to automatically identify Bragg peaks and separate them from diffuse backgrounds and analyze the crystalline lattice parameters for further analysis. A novel CNN architecture has been developed to identify anomalous background from detector instrumentation for dynamical cleaning of measurements. The implementation of automatic peak identification into IndeX allows scientists to visualize and analyze data in real-time. These methods can be applied in real-time during detector operation to improve experimental operations and scientific analysis. Our methods include a robust comparison with current analysis techniques which show improvement in a variety of aspects. These improvements will be incorporated into IndeX for visualization to allow scientists an interactive tool for crystal analysis.

Primary Keyword:

MLOps

Secondary Keyword:

anomaly detection

Tertiary Keyword:

ML-based optimization

Tools for Humans / 18

Computer vision for laboratory assistance tools**Authors:** Bianca Veglia¹; Francois Lemery¹; Frank Mayet¹; Ilya Agapov¹; Klaus Floettmann¹¹ DESY**Corresponding Authors:** ilya.agapov@desy.de, bianca.veglia@desy.de

This study focuses on the use of computer vision algorithms to improve the efficiency of laboratory tasks, data collection, and process monitoring. By deploying computer vision systems in the laboratory environment, various tasks can be automated and streamlined. This includes recognizing and tracking hardware, verifying positioning and providing AR modelling.

Here we present a project aimed at assisting laser laboratory operations by locating and uniquely identify equipment while tracking the relative position, as well as providing real-time assistance by accessing parameter settings and relevant information. The program combines object detection with optical character recognition and space mapping techniques. The real-time object classification is performed using the YOLO (You Only Look Once) model [1], a single-shot detector based on a Convolutional Neural Network backbone to form image features.

This contribution also discusses the challenges and potential limitations associated with implementing computer vision systems in laboratories, such as hardware requirements, data management concerns, and the need for machine learning models fine-tuned to the specific laboratory environment. In conclusion, the integration of computer vision techniques into laboratories represents a promising step towards more efficient and capable laboratory assistants.

[1] Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0). <https://github.com/ultralytics/ultralytics>

Primary Keyword:**Secondary Keyword:****Tertiary Keyword:**

Tools for Humans / 6

Towards Unlocking Insights from Logbooks using AI at DESY and BESSY**Authors:** Antonin Sulc¹; Gregor Hartmann²; Tim Wilksen¹¹ DESY MCS² BESSY**Corresponding Authors:** antonin.sulc@desy.de, gregor.hartmann@desy.de, tim.wilksen@desy.de

Logbooks store important knowledge of activities and events that occur during accelerator operations. However, orientation and automation of accelerator logbooks can be challenging due to various challenges like very technical texts or content being hidden in images instead of text. As AI technologies like natural language processing continue to mature, they present opportunities to address these challenges in the context of particle accelerator logbooks.

In this work, we explore the potential use of state-of-the-art AI techniques in particle accelerator logbooks. Our goals are to help operators increase the FAIR-ness (findability, accessibility, interoperability, reusability) of logbooks by exploiting the multimodal information to make everyday use easier with multimodal LLM (large language models).

Primary Keyword:

foundation models

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 95

Flash Talks for Poster/Demos 4th ICFA MaLAPA

Corresponding Authors: stjohnd@fnal.gov, miceli@fnal.gov

Compiled slide show for this year's poster/demo contributions.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 65

Integration of Multi-Objective Genetic Algorithm and neural networks in linac optimization.

Authors: Chanmi Kim¹; Chong Shik Park²

Co-author: eun-san Kim³

¹ *korea univ.*

² *Korea University, Sejong*

³ *Korea Univ. (Sejong)*

Corresponding Authors: kuphy@korea.ac.kr, physicexper@korea.ac.kr, eskim1@korea.ac.kr

Multi-Objective Genetic Algorithm(MOGA) is one of promising approach for optimizing nonlinear beam dynamics in accelerators. For explorative problems that have many variables and local optima, however the performance of MOGA is not always satisfactory. To improve the efficiency of optimization in linac beam line, we propose a novel integration of MOGA and neural networks. The neural network is trained with the data produced in the evolution of the MOGA. The objective values of the offspring are estimated with the trained neural network. Based on the estimated results, those offspring are ranked with the nondominated sorting method. We therefore propose a novel Machine Learning technique in which nonlinear tracking is replaced by two well-trained neural networks to beam line lattice.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 61

Simulation methods of 3D coupled storage ring based on SLIM formalism

Author: Jingyuan Zhao¹

Co-authors: Alexander Chao²; Chuanxiang Tang¹; Xiujie Deng¹; Zhilong Pan¹; Zizheng Li¹; liwei chen³

¹ Tsinghua University

² SLAC National Accelerator Laboratory & Tsinghua University

³ Tsinghua university

Corresponding Authors: 1400682808@qq.com, clw20@mails.tsinghua.edu.cn, li-zz18@mails.tsinghua.edu.cn, panzl@mail.tsinghua.edu.cn, tang.xuh@mail.tsinghua.edu.cn, dengxiujie@mail.tsinghua.edu.cn, achao@slac.stanford.edu

Over last decades, most synchrotron radiation light source designs are based on planar storage rings. Under the linear uncoupled condition, we can describe the physics of these storage rings using the auxiliary functions such as Twiss parameters. We can also give the effects of coupling on emittance by some approximations under the hypothesis of weak coupling. However, in recent years, the emittance of the synchrotron radiation light source storage ring has been reduced to the diffraction limit, the coupling effect can no longer be ignored. At the same time, some new light sources are also trying to use coupling effects to achieve higher goals. Therefore, it is very important to study and calculate coupling effects accurately and self-consistently without too many hypothesis. The SLIM formalism, developed by Prof Alexander Chao, can help solve the problem. SLIM is a linear storage ring beam dynamic formalism based on matrices. It can directly calculate the trajectory of the electron distribution center and the equilibrium beam size and shape in phase space without tracking or introducing auxiliary functions. It also naturally includes the coupling of horizontal, vertical and longitudinal motions in the result. It is a good tool for linear storage ring (with or without coupling) design. However, few researchers are currently using SLIM. We did some exploration of SLIM's application. We introduced the thick lens 7-dimensional matrix and the analytical solution of the contribution to the quantum diffusion rate from radiation elements into SLIM to avoid slicing the element and then improve the computational efficiency of the code. Then we tried to combine the speed-up SLIM with MOGA for a new coupling based storage ring design and optimization. In the future, the speed-up SLIM may provide a fast physical computing kernel or fast generation of datasets for machine learning based storage ring designs.

Primary Keyword:

foundation models

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

surrogate model architectures

Poster/Demos / 89

Online optimisations of lifetime and injection efficiency in the ESRF EBS storage ring

Author: Nicola Carmignani¹

Co-authors: Lee Carver¹; Lina Hoummi¹; Simon White¹; Simone Liuzzo¹; Thomas Perron¹

¹ ESRF

Corresponding Authors: thomas.perron@esrf.fr, lee.carver@esrf.fr, lina.hoummi@esrf.fr, nicola.carmignani@esrf.fr, simon.white@esrf.fr, simone.liuzzo@esrf.fr

Multi-objective optimisations are extensively used during the design phase of the modern storage rings, to optimise at the same time the simulated Touschek lifetime and injection efficiency. On-line optimisations of either the measured lifetime or the measured injection efficiency have been also extensively used at the ESRF and in several other accelerators. Online multi-objective optimisations are technically more complicated, but can also be performed. Different online optimisations performed at the ESRF EBS storage ring using different algorithms will be presented.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 17

Machine Learning Based Response Matrix Correction

Author: liwei chen¹

¹ *Tsinghua university*

Corresponding Author: clw20@mails.tsinghua.edu.cn

The response matrix is the closed orbit distortion at each BPM responses to the change in every corrector. For a large ring, the response matrix has tens of thousands of data points which can fully include the linear optics of the ring. LOCO use response matrix for lattice calibration and error correction. For 4th generation diffraction limitation ring which uses many strong sextupoles and octupoles, the response matrix will influence by the nonlinearity and can only be driven from closed orbit distortion tracking not from linear matrix. The strong nonlinearity will make it difficult for LOCO to match lattice parameters and also need more time to get jacobi matrix. Machine learning may help bypass the time assuming jacobi matrix and avoid local optima. This work try to improve the speed and accuracy of LOCO by machine learning method.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 32

Orbit correction by machine learning in TPS storage ring

Author: Mau-Sen Chiu¹

¹ *National Synchrotron Radiation Research Center*

Corresponding Author: chiu.ms@nsrrc.org.tw

Machine learning has been widely used in many fields including science and technology. This paper will focus on the orbit correction by the algorithm of neural networks, a subset of machine learning,

in Taiwan photon Photon Source. The training data for neural network is collected by machine study and accelerator toolbox (AT).

Primary Keyword:

AI-based controls

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 1

Comparing gradient-based and non-gradient modelling and optimisation methods for investigating synchrotron dynamics

Author: Seb Wilkes¹

¹ *University of Oxford*

Corresponding Author: juan.martinezwilkes@physics.ox.ac.uk

Differentiable modelling has garnered significant interest in the accelerator physics community, but literature is lacking on its specific application to synchrotron dynamics. In principle, access to the gradients should reduce the number of trials required in an optimisation loop. As a 'real test case', we want to optimise the best set of beam perturbations to achieve the goals of 'Pulse Picking By Resonant Excitation' – a mode of operation that caters to timing-mode users. By comparing the application of gradient-descent methods, using gradients computed from JAX, to gradient-free methods, I discuss the applicability of each approach.

Primary Keyword:

differentiable models

Secondary Keyword:

bayesian optimization

Tertiary Keyword:

Poster/Demos / 11

Online image-based beam-dump anomaly detection

Author: Francisco Moreira Huhn^{None}

Co-author: Francesco Velotti¹

¹ *CERN*

Corresponding Authors: francesco.maria.velotti@cern.ch, francisco.huhn@cern.ch

The CERN SPS Beam Dump System (SBDS) disposes the beam in the SPS at end of cycled operation or in case of machine malfunctioning, with its kicker magnets deviating the beam to an absorber block and diluting the particle density. This is a critical system, as its malfunctioning can lead to absorber block degradation, unwanted activation of the surroundings or even damage to the vacuum chamber. We develop an online anomaly detection system for the SBDS based on real-time images of a beam screen device. Crucially, the model must accurately classify these images despite being trained on

an unlabelled dataset and one in which anomalous samples are uncommon. We show this can be achieved with a convolutional autoencoder and by leveraging the quality of its reconstructions. This work improves the safety of operating the SPS and contributes towards the goal of automating the operation of particle accelerators.

Primary Keyword:

anomaly detection

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 72

Multiobjective Optimization of Cyclotron Cavity Model using Neural Network

Author: Ahsani Hafizhu Shali¹

Co-authors: Hiroki Kanda ¹; Mitsuhiro Fukuda ¹; Takafumi Hara ¹; Tetsuhiko Yorita ¹

¹ *Research Center for Nuclear Physics, Osaka University*

Corresponding Authors: takafumi@rcnp.osaka-u.ac.jp, ahsahafi@rcnp.osaka-u.ac.jp, kandrah@rcnp.osaka-u.ac.jp, yorita@rcnp.osaka-u.ac.jp, mhufukuda@rcnp.osaka-u.ac.jp

Design optimization for a cyclotron is important for obtaining a high accelerating voltage to increase efficiency. A different cavity geometry made from the same material usually will have a different quality factor, which might affect the turn separation, especially if the electric field at the accelerator zones changes. For this purpose, a neural network is trained to give predictions of several accelerating cavity quantities given some initial parameters related to the geometry of the cavity, using training data obtained from an electromagnetic numerical solver for rf cavity. The optimization is done using the multiobjective optimization scheme due to the fact that optimization is constrained by several operational parameters, such as the RF frequency. We show that the use of neural network combined with multiobjective optimization can be implemented for cyclotron cavity design optimization.

Primary Keyword:

ML-based optimization

Secondary Keyword:

surrogate model architecture

Tertiary Keyword:

Poster/Demos / 27

RL-Based Control Strategies for HIPI Accelerator

Author: Chunguang Su¹

Co-author: Xiaolong Chen ²

¹ *Institute of modern physics, Chinese Academy of Sciences*

² *Institute of modern physics*

Corresponding Authors: suchunguang@impcas.ac.cn, chenxiaolong@impcas.ac.cn

Beam commissioning is a key procedure to achieve high quality beam. Conventional “Monkey ump” tuning is time-consuming and inefficient. Reinforcement learning (RL) can swiftly make decisions based on the current system state and control requirements, providing an efficient control solution for accelerator systems.

High Intensity Proton Injector (HIPI) accelerator requires a rapid and effective control method to meet user demands. To attain this aim, initially, a neural network-based surrogate model is created by collecting HIPI operational data. Subsequently, an RL-based strategy, based on the surrogated model, is used to control the components of HIPI after training on different initial states. Finally, the policy undergoes ten rounds of validation on HIPI. The results consistently illustrate the strategy’s capacity to improve beam transmission efficiency within minutes, showcasing the potential of RL in solving particle accelerator control challenges.

Primary Keyword:

AI-based controls

Tertiary Keyword:

Secondary Keyword:

reinforcement learning

Poster/Demos / 39

Reinforcement Learning Based Radiation Optimization at a Linear Accelerator

Author: Chenran Xu¹

Co-authors: Andrea Santamaria Garcia ²; Anke-Susanne Müller ³

¹ *Karlsruhe Institut für Technologie (KIT)*

² *Karlsruhe Institute of Technology*

³ *KIT*

Corresponding Authors: andrea.santamaria@kit.edu, chenran.xu@kit.edu, anke-susanne.mueller@kit.edu

Low energy linear accelerators can generate intense ultra-short THz pulses of coherent synchrotron radiation (CSR) by using chicanes and/or undulators to bend the path of the electron bunch. Additionally, potential users of the THz light might have particular requests for their experiments, which calls for a way to more flexibly tailor the emitted spectrum.

It’s often a complex and time-consuming task to optimize the accelerator setting for maximal radiation outcome, as the input parameters are often correlated and the system response is non-linear.

In this contribution, we apply reinforcement learning techniques to optimize the linear accelerator FLUTE at KIT, with the goal to maximize its THz pulse generation. The agent is trained in a high-speed simplified simulation model. The utilization of domain randomization allows the pre-trained RL agent to generalize its policy to higher-fidelity simulations and different accelerator setups, indicating its potential for real-world tasks.

Primary Keyword:

ML-based optimization

Secondary Keyword:

reinforcement learning

Tertiary Keyword:

Poster/Demos / 33

Dynamic vacuum and temperature predictions for informed anomaly detection at the CERN-SPS

Author: Francesco Velotti¹

Co-authors: Mike Barnes¹; Verena Kain¹

¹ CERN

Corresponding Authors: mike.barnes@cern.ch, verena.kain@cern.ch, francesco.maria.velotti@cern.ch

High intensity beams, through electron cloud and impedance based mechanisms, cause an increase in vacuum pressure in the SPS kicker magnets. These magnets are pulsed at high voltage in order to quickly deflect the beam. However, if the vacuum inside their aperture deteriorates, it can lead to an electrical breakdown and potential damage to the kicker itself.

Conversely, a breakdown in a kicker that is not initiated by beam behaviour exhibits a very similar signature - a rapid increase in pressure. Additionally, both an electrical breakdown and rapid increase in dynamic vacuum can occur simultaneously, making it challenging to detect sparks in the magnet and automate the process. The rarity of breakdown events further complicates matters, as simple classification of vacuum signatures cannot be applied.

To address this issue, we propose a data-driven model that predicts vacuum and temperature behavior and distinguishes between normal and anomalous activities.

Primary Keyword:

failure prediction

Secondary Keyword:

anomaly detection

Tertiary Keyword:

Poster/Demos / 54

Trust Region Bayesian Optimization for Online Accelerator Control

Authors: Auralee Edelen¹; Nikita Kuklev²; Ryan Roussel¹; Simone Liuzzo³; Tobias Boltz¹

¹ SLAC National Accelerator Laboratory

² Argonne National Laboratory

³ ESRF

Corresponding Authors: tboltz@slac.stanford.edu, nkuklev@anl.gov, simone.liuzzo@esrf.fr, edelen@slac.stanford.edu, rroussel@slac.stanford.edu

Bayesian optimization (BO) is an effective tool in performing online control of particle accelerators. However, BO algorithms can struggle in high dimensional or tightly constrained parameter spaces due to its inherent bias towards over-exploration, leading to slow convergence times for relatively simple problems and high likelihoods of constraint violations. In this work, we describe the application of Trust Region BO (TurBO), which dynamically limits the size of parameter space when performing optimization or characterization. As a result, the convergence of BO towards local extrema is greatly enhanced and the number of constraint violations is substantially reduced. We describe the performance of TurBO on benchmark problems as well as in experiments at accelerator facilities including ESRF and AWA.

Primary Keyword:

bayesian optimization

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 20

Anomaly Detection for Diode Failures

Authors: Jennefer Maldonado¹; Jonathan Laster¹

Co-authors: Donald Bruno ¹; Seth Nemesure ¹

¹ *Brookhaven National Laboratory*

Corresponding Authors: bruno@bnl.gov, seth@bnl.gov, jlaster@bnl.gov, jmaldonad@bnl.gov

The Collider-Accelerator Department's (C-AD) Controls Group at Brookhaven National Laboratory produces and implements tools to analyze data after magnet quench events. Diodes are used in the circuitry to protect quenching magnets from damage. Intermittently failing diodes can be difficult to identify as they may not always impact beam. Accelerator physicists have studied the voltage tap shutoff curves at various times after a failure, identifying specific time zones over which a derivative may be calculated to detect an anomaly. Anomaly detection and clustering models show promise by detecting negative events and outliers in datasets. Using machine learning modeling algorithms, an automated analysis for each power supply based on the voltage tap data can be applied which will help to efficiently identify faulty diodes and limit the number of false positives reported. This could potentially lead to faster recovery times as well as help avoid equipment damage.

Primary Keyword:

anomaly detection

Secondary Keyword:

failure prediction

Tertiary Keyword:

AI-based controls

Poster/Demos / 14

Multi objective Bayesian optimization of ECR ion source at the Linear IFMIF Prototype Accelerator.

Author: Andrea De Franco¹

Co-authors: Tomoya Akagi ¹; Benoit Bolzon ²; Nicolas Chauvin ²; Fabio Cismonti ³; Herve Dzitko ³; Tomonobu Itagaki ¹; Antti Jokinen ³; Keitaro Kondo ¹; Kai Masuda ¹; Takahide Nakayama ¹; Masayoshi Sugimoto ¹

¹ *National Institutes for Quantum Science and Technology (QST)*

² *French Alternative Energies and Atomic Energy Commission (CEA)*

³ *Fusion for Energy (F4E)*

Corresponding Author: defranco.andrea@qst.go.jp

The Linear IFMIF Prototype Accelerator (LIPAc) is designed to validate the main key technical solutions for the particle accelerator of the International Fusion Materials Irradiation Facility, which will answer the need of the fusion community for a high energy (14.1 MeV) high intensity neutron source. LIPAc is jointly developed under the EU-Japan Broader Approach agreement to accelerate 125mA of D+ to 9MeV in Continuous Wave. Its Electron Cyclotron Resonance (ECR) ion source was developed by CEA and successfully demonstrated >150mA D+ beam at 100 keV in CW. In this work we present the multi objective Bayesian optimization of several of its tunable parameter (confinement magnetic field, RF power, gas flow, etc.) to find the most suitable compromise between average extracted beam current and its stability over time.

Primary Keyword:

bayesian optimization

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

Poster/Demos / 67

Optimization of a Longitudinal Bunch Merge Gymnastic with Reinforcement Learning

Author: Yuan Gao^{None}

Co-authors: Kevin Brown ; Auralee Edelen ; Georg Hoffstaetter ; Weijian Lin ; John Morris ; David Sagan ; Vincent Schoefer ; Malachi Schram ; Yinan Wang ; Keith Zeno ; Michael Costanzo ; Jennefer Maldonado ¹; Jonathan Unger ; Eiad Hamwi

¹ Brookhaven National Laboratory

Corresponding Authors: wangy88@rpi.edu, ygao@bnl.gov, brownk@bnl.gov, schoefer@bnl.gov, georg.hoffstaetter@cornell.edu, wl674@cornell.edu, jmaldonad@bnl.gov, jeu8@cornell.edu, zeno@bnl.gov, jtm@bnl.gov, schram@jlab.org, david.sagan@cornell.edu, edelen@slac.stanford.edu

The RHIC heavy ion program relies on a series of RF bunch merge gymnastics to combine individual source pulses into bunches of suitable intensity. Intensity and emittance preservation during these gymnastics require careful setup of the voltages and phases of RF cavities operating at several different harmonic numbers. The optimum setting tends to drift over time, degrading performance and requiring operator attention to correct. We describe a reinforcement learning approach to learning and maintaining an optimum configuration, accounting for the relevant RF parameters and external perturbations (e.g., a changing main dipole field) using a physics-based simulator at BNL Booster.

Primary Keyword:

ML-based optimization

Secondary Keyword:

reinforcement learning

Tertiary Keyword:

AI-based controls

Poster/Demos / 10

Distance Preserving Machine Learning for Uncertainty Aware Accelerator Capacitance Predictions

Author: Steven Goldenberg¹

Co-authors: Chris Pappas²; Dan Lu²; Jared Walden²; Kishansingh Rajput¹; Majdi I. Radaideh³; Malachi Schram¹; Sarah Cousineau²; Sudarshan Harave⁴; Thomas Britton¹

¹ *Thomas Jefferson National Accelerator Facility*

² *Oak Ridge National Laboratory*

³ *University of Michigan*

⁴ *SLAC National Accelerator Laboratory*

Corresponding Authors: lud1@ornl.gov, haraves@slac.stanford.edu, scousine@ornl.gov, waldenjd@ornl.gov, sgolden@jlab.org, pappasgc@ornl.gov, tbritton@jlab.org, radaideh@umich.edu, kishan@jlab.org, schram@jlab.org

Providing accurate uncertainty estimations is essential for producing reliable machine learning models, especially in safety-critical applications such as accelerator systems. Gaussian process models are generally regarded as the gold standard method for this task, but they can struggle with large, high-dimensional datasets. Combining deep neural networks with Gaussian process approximation techniques has shown promising results, but dimensionality reduction through standard deep neural network layers is not guaranteed to maintain the distance information necessary for Gaussian process models. We build on previous work by comparing the use of the singular value decomposition against a spectral-normalized dense layer as a feature extractor for a deep neural Gaussian process approximation model and apply it to a capacitance prediction problem for the High Voltage Converter Modulators in the Oak Ridge Spallation Neutron Source. Our model shows improved distance preservation and predicts in-distribution capacitance values with less than 1% error.

Primary Keyword:

uncertainty quantification for ML

Secondary Keyword:

failure prediction

Tertiary Keyword:

Poster/Demos / 40

Improving Surrogate Model Performance for Sparse Outputs in the Spatial Domain

Authors: Jaehoon Cha¹; Kathryn Baker²; Susmita Basak¹

¹ *STFC / Scientific Machine Learning*

² *STFC / ISIS*

Corresponding Authors: jaehoon.cha@stfc.ac.uk, susmita.basak@stfc.ac.uk, k.baker@stfc.ac.uk

Neural networks have a strong history as universal function approximators and as such have seen extensive use as surrogate models for computationally expensive physics simulations. However, achieving predictions for sparse or step functions are difficult, particularly within the spatial domain. This is of particular concern when these sparse events in the spatial dimension are used to calculate other outputs as the cumulative effect of the error in the peak reconstruction would result in vastly different results to that of the ground truth. This is a common case in surrogate models for particle

accelerators, when machine settings are used to predict beam descriptors along the spatial dimension of some accelerator component.

This work will discuss two solutions for mitigating the problem of step function reproducibility for neural networks. Firstly, by introducing a Fourier feature transformation and secondly by using a loss function that enforces co-dependence between outputs during training. It considers as a case study a surrogate model developed to predict key beam descriptors in the Medium Energy Beam Transport (MEBT) at the ISIS Neutron and Muon Source.

Primary Keyword:

surrogate model architecture

Secondary Keyword:

surrogate model tuning

Tertiary Keyword:

digital twins

Poster/Demos / 31

Optimization design of photocathode injector assisted by deep Gaussian process

Authors: Zheng Sun¹; Xin Tianmu²

¹ *Institution of high energy physics,IHEP*

² *IHEP*

Corresponding Authors: sunzheng@ihep.ac.cn, xintm@ihep.ac.cn

To meet the requirement of electron beam characteristics at linac entrance of CEPC and PWFA. A method of searching in a high-dimensional parameter space was performed using a multi-objective genetic algorithm. A deep Gaussian process was adopted as an surrogate model to solve high-dimensional parameter optimization problem. geometric parameters of radio frequency gun and beam element parameters have been determined to minimize rms transverse beam emittance and rms bunch length. In conclusion, we conducted a study on bunch injectors with three distinct charge levels. We provided optimization results for initial charge levels of 1nC, 5nC, and 10nC, corresponding to flat-top beam distributions and Gaussian beam distributions. These injectors were composed of an L-band RF gun, a pair of solenoids, and an accelerator tube. Our findings demonstrate that the optimization algorithm enables us to efficiently identify multiple sets of optimal solutions. Furthermore, for high-dimensional parameter optimization, the utilization of a deep Gaussian process has proven to be a favorable method.

Primary Keyword:

ML-based optimization

Secondary Keyword:

surrogate model architecture

Tertiary Keyword:

bayesian optimization

Poster/Demos / 56

Development of beam transport system optimization method using VAE and Bayesian optimization

Author: Yasuyuki Morita¹

Co-authors: Takahiro Nishi¹; Takashi Nagatomo¹; Takashi Washio²; Yuta Nakashima³

¹ *RIKEN*

² *The Institute of Scientific and Industrial Research, Osaka University*

³ *IDS, Osaka University*

Corresponding Authors: nagatomo@riken.jp, takahiro.nishi@riken.jp, washio@ar.sanken.osaka-u.ac.jp, n-yuta@ids.osaka-u.ac.jp, morita16@rcnp.osaka-u.ac.jp

In recent years, Bayesian optimization has been attracting attention as a tuning method for accelerators. However, the number of iterations required increases as the number of parameters increases. Therefore, there is a limit to the number of parameters that can be optimized in a realistic amount of time. In this study, we proposed a new method that combines the dimensionality reduction method with VAE. This method increases the number of parameters that can be handled at one time and enables the tuning of long beam transport systems in a short time. In this presentation, we will report on the preparation status of the demonstration experiment.

Primary Keyword:

AI-based controls

Secondary Keyword:

bayesian optimization

Tertiary Keyword:

Poster/Demos / 36

Experience with ML-driven applications at PETRA III

Authors: Bianca Veglia¹; Ilya Agapov¹

¹ *DESY*

Corresponding Authors: bianca.veglia@desy.de, ilya.agapov@desy.de

We present experience with deploying several ML-based methods for control and optimization of the PETRA III storage ring.

First, we discuss the recent progress with the compensation of influence of insertion devices (IDs) on beam orbit using Deep Neural Networks. Different models were trained to predict the distortion in the closed orbit induced by movements of the IDs in the context of a feed-forward correction scheme. The networks accurately predict the transverse displacement of the beam along the ring for any given ID configuration allowing to compute the appropriate compensation.

In another activity, Bayesian Optimization routines implemented in the Badger/Xopt [1,2] optimization framework were used interfacing with the machine control system. Optimizations of lifetime, and injection efficiency were performed and compared with the results obtained with the current manual procedures during machine dedicated times.

Finally, the experience with the ACSS (Accelerator Control and Simulation Services) pipeline [3] is described together with application examples on PETRA III. The software framework provides an environment for scheduling and orchestrating of multiple intelligent agents, training and tuning of ML models, handling of data streams and for software testing and verification, addressing the need

for significant reduction of the amount of human intervention in AI-based operation. The pipeline was successfully tested using the orbit correction service on the Petra III control network.

- [1] Zhang, Z., et al., “Badger: The Missing Optimizer in ACR”, IPAC’22
- [2] Roussel, R., et al., “Xopt: A Simplified Framework for Optimization of Accelerator Problems Using Advanced Algorithms”, IPAC’23
- [3] Agapov, I., Böse, M., Malina, L., “A Pipeline for Orchestrating Machine Learning and Controls Applications”, IPAC’22

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 76

Rapid Tuning or Synchrotron Surrogate Model at the Recycler Ring

Author: Jason St. John¹

¹ *Fermilab*

Corresponding Author: stjoh@fnal.gov

The 8 GeV proton-storage Recycler Ring (RR) is essential for reaching megawatt beam intensity goals for the DUNE neutrino beam at Fermilab. Custom shims on each RR permanent magnet were designed to cancel manufacturing defects and bring magnetic fields to the design values. Remaining imperfections cause the observed tune variation vs energy to deviate from what is calculated using the design fields. Using the POUNDERS (“Practical Optimization Using No Derivatives for sums of Squares”) optimization method with Synergia in the loop, we demonstrate rapid convergence to a set of additive, higher-order multipole moments of these magnetic shims which reproduce that observed variation, and show that the convergence advantage grows with the parameter-space dimensionality.

Primary Keyword:

ML-based optimization

Secondary Keyword:

surrogate model tuning

Tertiary Keyword:

digital twins

Poster/Demos / 77

Analysis and Improvement of Generalisability of Anomaly Detection Methods

Authors: Kathryn Baker¹; Mateusz Leputa²; Mihnea Romanovschi³

¹ *STFC / ISIS*

² UKRI-STFC-ISIS³ Science and Technology Facilities Council (STFC) UKRI ISIS Neutron and Muon Source**Corresponding Authors:** mihnea.romanovschi@stfc.ac.uk, k.baker@stfc.ac.uk

Machine Learning (ML) has gained significant prominence in the field of engineering due to its adaptability and versatility. An example of its practical application is in anomaly detection, which serves the fundamental purpose of providing a binary response to the question, “Has an issue arisen?”. Most machine learning and anomaly detection projects strive to provide generalisable solutions that are robust to changes within systems. This is of particular importance when components that are subject to anomalies are upgraded; the behaviour and values in the system remain unchanged, but the data resolution may be higher, and anomalies may manifest themselves in slightly different ways. An ideal ML solution would be able to detect anomalies as accurately in this new system as in the old one with minimal intervention or retraining. However, cases like the one described do not happen frequently and therefore it is difficult to test generalisability of models. The methane moderator of Target Station 1 at ISIS originally had an anomaly detection model. Following an upgrade, this work will explore how well the original model generalizes for the system such that it can be easily adapted from the old version of the moderator to the newer one. The original model used a combination of a 1 dimensional convolutional neural network binary classifier and a hypothesis test. The outputs of those two models would be used in a weighted sum. We will also investigate other methods to improve the generalizability that will allow for more flexibility with minimal changes or training when adapting the model from an old training set to a newer one.

Primary Keyword:

anomaly detection

Secondary Keyword:**Tertiary Keyword:****Poster/Demos / 48**

The Reinforcement Learning for Autonomous Accelerators International Collaboration

Authors: Andrea Santamaria Garcia¹; Chenran Xu²; Jan Kaiser³; Simon Hirländer⁴**Co-author:** Annika Eichler⁵¹ Karlsruhe Institute of Technology² Karlsruhe Institut für Technologie (KIT)³ Deutsches Elektronen-Synchrotron DESY⁴ University of Salzburg⁵ Deutsches Elektronen Synchrotron DESY**Corresponding Authors:** jan.kaiser@desy.de, simon.hirlander@plus.ac.at, andrea.santamaria@kit.edu, chenran.xu@kit.edu

Reinforcement Learning (RL) is a unique learning paradigm that is particularly well-suited to tackle complex control tasks, can deal with delayed consequences, and learns from experience without an explicit model of the dynamics of the problem. These properties make RL methods extremely promising for applications in particle accelerators, where the dynamically evolving conditions of both the particle beam and the accelerator systems must be constantly considered.

While the time to work on RL is now particularly favourable thanks to the availability of high-level programming libraries and resources, its implementation in particle accelerators is not trivial and requires further consideration.

In this context, the Reinforcement Learning for Autonomous Accelerators (RL4AA) international collaboration was established to consolidate existing knowledge, share experiences and ideas, and collaborate on accelerator-specific solutions that leverage recent advances in RL.

The collaboration was launched in February 2023 during the RL4AA'23 workshop at the Karlsruhe Institute of Technology, and the second workshop is held in Salzburg, Austria in February 2024. These workshops feature keynote lectures by experts, technical presentations, advanced tutorials, poster sessions, and contributions on RL applications in various facilities. The next upcoming workshop will be held in February 2023 at DESY, Hamburg.

Primary Keyword:

reinforcement learning

Secondary Keyword:

AI-based controls

Tertiary Keyword:

Poster/Demos / 94

Simultaneous corrections of nonlinear errors in the LHC triplets using machine learning

Author: Felix Carlier¹

Co-author: Elena Fol¹

¹ CERN

Corresponding Authors: elena.fol@cern.ch, felix.sion.carlier@cern.ch

Non-linear optics commissioning for the LHC has faced challenges with higher order errors using a diverse array of correction techniques. Feed down of these errors complicates the correction process, demanding significant time and effort. As machine complexity increases and IP beta functions decrease, there is a growing need for efficient and reliable correction methods. This study explores the use of new machine learning methods to simultaneously correct errors of multiple orders. Leveraging MAD-NGs computation speeds presents great promise in the realm of machine learning for optics. Results from simulations using these novel methods are presented and show significant improvements compared to classical approaches currently used.

Primary Keyword:

uncertainty quantification for ML

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

Poster/Demos / 86

Application of Machine Learning to Accelerator Operations at SACLASPring-8.

Author: Eito Iwai¹

Co-authors: Hirokazu Maesaka²; Ichiro Inoue³

¹ SPring-8

² RIKEN/JASRI

³ *RIKEN***Corresponding Authors:** iwai@spring8.or.jp, maesaka@spring8.or.jp, inoue@spring8.or.jp

We've introduced Machine Learning methods to accelerator operations at SACLA/SPRING-8. One of them is an automatic beam tuning based on Bayesian Optimization. In the initial test, we tried to maximize the pulse energy by using the optimizer. Then we've introduced a new high-resolution single-shot inline spectrometer (resolution of a few eV) to maximize the spectral brightness. Today the optimizer is applied for various beam tuning. Another activity is to Anomaly Detection of Thyratrons. Based on the rate of misfiring and its grid waveform, Failure Prediction of working thyratrons are evaluated. These ML related activities and their status are reported.

Primary Keyword:

bayesian optimization

Secondary Keyword:

failure prediction

Tertiary Keyword:**Poster/Demos / 92****End-to-end Simulations and ML infrastructure for Light Sources****Author:** Jonathan Edelen¹¹ *RadiaSoft LLC***Corresponding Author:** jedelen@radiasoft.net

SLAC and RadiaSoft are partnering to provide integration support for two parallel workflows that support end-to-end modeling and machine learning integration for accelerators. LUME, Light Source Unified Modeling Environment, has been developed by SLAC to facilitate end-to-end modeling for machine tuning and optimization. This workflow includes the integration of machine learning surrogate models. In parallel, RadiaSoft is developing a workflow, Omega, for chaining simulations developed using different Sirepo applications. This tool allows users to import simulations built in Sirepo and connect them into an end-to-end simulation. Our collaboration is focused on integrating these two workflows in order to provide the community with a unified toolbox for online modeling of light sources. This tutorial will provide a hands-on example for both Lume and Omega that showcases their interoperability.

Primary Keyword:

MLOps

Secondary Keyword:

AI-based controls

Tertiary Keyword:**Poster/Demos / 90**

Improving Accelerator Surrogate Models with a Knowledge of Physics

Author: Christopher Pierce¹

Co-author: Young-Kee Kim ¹

¹ *University of Chicago*

Corresponding Authors: chris@chris-pierce.com, ykkim@hep.uchicago.edu

Physics-based simulation tools are essential to the design and operation of modern particle accelerators. Although accurate, these tools tend to be expensive to evaluate, aren't always compatible with modern implementations of automatic differentiation, and have a hard time incorporating data from real-world machines. Surrogate models have the potential to solve these problems by turning simulated data (or measurements) into quick-to-evaluate and differentiable functions through machine learning techniques. In this work, we improve on the way surrogate models are trained by “teaching” them about the already known laws of physics that govern charged particle beams. By training models of a toy particle accelerator with an additional “physics-informed neural network” loss term based on the Vlasov equation, we demonstrate significantly lower error (by more than a factor of two) in our models compared to those trained on simulated data alone. This research paves the way for accelerator physics to bridge the gap between purely data-driven models and physics-based simulation tools by introducing physics-informed priors for surrogate modeling.

Primary Keyword:

digital twins

Secondary Keyword:

surrogate model tuning

Tertiary Keyword:

surrogate model architectures

Poster/Demos / 15

Reshaping SRF Cavity Resonance Management with Smart Techniques

Author: Faya Wnag¹

¹ *SLAC*

Corresponding Author: fywang@slac.stanford.edu

Motion control is assuming an increasingly pivotal role within modern large accelerator facilities, such as 4th generation storage ring-based light sources, SRF accelerators, and high-performance photon beamlines. For very high-Q SRF linacs, such as LCLS-II, the precise management of cavity resonance becomes indispensable for maintaining stable operations. Failing to do so would entail a significant upsurge in RF power requirements, consequently increasing operational and capital costs due to the necessity for additional RF power sources. We have developed an intelligent cavity resonance controller founded on a data-driven model, featuring an exceptionally lightweight surrogate model engineered to address the intricate dynamics of forced cavities in the presence of microphonics and coupled with nonlinear Lorentz forces. The effectiveness of this model has been rigorously validated through real SRF cavities at SLAC. We are currently in the process of implementing the controller on hardware, specifically the existing LLRF system of LCLSII. Building on the success of this work, the model can be expanded to encompass general motion controls where exceptionally low-tolerance vibration is required. In this presentation, we will introduce the model and provide an overview of the latest test results.

Primary Keyword:

AI-based controls

Secondary Keyword:

ML-based optimization

Tertiary Keyword:

active learning

Poster/Demos / 50

Reinforcement Learning for Intensity Tuning at Large FEL Facilities

Authors: Jan Kaiser¹; Annika Eichler²; Auralee Edelen³; Daniel Ratner³

¹ *Deutsches Elektronen-Synchrotron DESY*

² *DESY*

³ *SLAC National Accelerator Laboratory*

Corresponding Authors: annika.eichler@desy.de, edelen@slac.stanford.edu, jan.kaiser@desy.de

One of the key metrics determining the capabilities of Free Electron Laser (FEL) facilities is the intensity of photon beam they can provide to experiments. However, in day-to-day operations, tuning to maximise the FEL intensity is one of the most difficult and time-consuming tasks. Skilled human operators still need large amounts of the available beam time, which are then not available for experiments, to achieve maximum performance. The large number of tuning parameters and high non-linearity of the underlying dynamics have so far made it challenging to develop autonomous FEL tuning solutions. We present a method based on reinforcement learning to train a neural network policy to autonomously tune the FEL intensity at *LCLS* and *European XFEL*. Our method is trained requiring little to no beam time and is appealing for tuning across different FEL setups. In contrast to conventional black box optimisation approaches that do not share information across different tuning sessions and setups, a trained policy can leverage its experience to tune the FEL intensity with minimal online exploration.

Primary Keyword:

reinforcement learning

Secondary Keyword:

AI-based controls

Tertiary Keyword:

ML-based optimization

Poster/Demos / 73

Research on Recognition of Quench and Flux Jump Based on Machine Learning

Author: Bao Niu¹

¹ *Institute of Modern Physics, Chinese Academy of Sciences*

Corresponding Author: niubao@impcas.ac.cn

The Institute of Modern Physics is developing the Fourth generation of Electron Cyclotron Resonance (FECR), which requires Nb3Sn superconducting hexapole magnets with higher magnetic fields and composite structures. For Nb3Sn superconducting magnets, they exhibit significant thermal magnetic instability, known as “flux jump”. This characteristic can generate random voltage spikes during the excitation process of the magnet, leading to misjudgment of the Quench Detection System (QDS) and seriously affecting the normal operation of FECR. To solve this problem, this study uses machine learning algorithms and aims to build a simplified and efficient recognition model to effectively distinguish the phenomenon of overshoot and flux jump during the excitation process of Nb3Sn magnets. Based on the voltage data obtained from multiple excitation processes of Nb3Sn superconducting hexapole magnets, this paper extracted 27 quench samples and 25 flux jump samples, and extracted 33 features from each sample. Multiple machine learning algorithms were used to train and construct these data, and the accuracy of different algorithms was compared to ultimately explore the best recognition model. The experimental results show that the model only uses 5 features and achieves 100% classification accuracy on linear kernel SVM. By using this machine learning model, high accuracy and computational speed have been achieved in the recognition of magnetic flux jump and quench, which can provide reference for the optimization of subsequent FECR quench detection algorithms.

Primary Keyword:

anomaly detection

Secondary Keyword:

foundation models

Tertiary Keyword:

Poster/Demos / 98

Towards Natural Language-driven Autonomous Particle Accelerator Tuning

Authors: Anne Lauscher¹; Annika Eichler²; Jan Kaiser³

¹ *University of Hamburg*

² *Deutsches Elektronen Synchrotron DESY*

³ *Deutsches Elektronen-Synchrotron DESY*

Corresponding Author: jan.kaiser@desy.de

Autonomous tuning of particle accelerators is an active and challenging field of research with the goals of reducing tuning times and enabling novel accelerator technologies for novel applications. Large language models (LLMs) have recently made enormous strides towards the goal of general intelligence, demonstrating that they are capable of solving complex task based just a natural language prompt. Here we demonstrate how LLMs can be used for autonomous tuning of particle accelerators using natural language. We test our approach on commonly performed tuning task at the ARES accelerator facility at DESY, and briefly compare its performance to other state-of-the-art autonomous accelerator tuning methods. Ultimately, this line of work could enable operators of particle accelerators to request working points through natural language and collaborate with autonomous tuning algorithms in an intuitive way, thereby significantly simplifying the operation of these complex and high-impact scientific facilities

Primary Keyword:

AI-based controls

Secondary Keyword:

Tertiary Keyword:

Poster/Demos / 101

Bayesian Optimal Experimental Design for AGS Booster Magnet Misalignment Estimation

Author: Weijian Lin¹

¹ *Cornell University*

Corresponding Author: wl674@cornell.edu

Magnet control is important to improve beam quality as its misalignments cause beam degradation and prevent the beam from reaching the desired specifications (e.g., polarization). Magnet misalignment measurements serve as the reference values in operations and provide a foundation for effective control. However, use of the historical measurement data may cause a significant deviation from the target physical system, as the magnets shift over time. Due to a lack of accurate reference values, the current practice of beam control relies mainly on empirical tuning by experienced operators, which may be inefficient or sub-optimal. To address this, we propose a Bayesian optimal experimental design (BOED)-based approach for identifying the magnet misalignments using a Bmad model of the Booster—one of the synchrotrons in the RHIC pre-accelerator chain. In the present application, the BOED approach is used to determine magnet control variables (i.e., currents) which are expected to lead to beam position data that most reduces uncertainty in the magnet misalignment parameters. This data is then used to calibrate the physical model of the Booster, leading to a more accurate simulation model for future polarization optimizations. This case study represents a new calibration paradigm for accelerator operations that makes use of models and data to optimally guide experiments and, ultimately, improve polarization performance in RHIC via uncertainty reduction.

Primary Keyword:

Secondary Keyword:

Tertiary Keyword:

Tools for Humans / 21

Improving Electronic Logbook Searches Using Natural Language Processing

Author: Jennefer Maldonado¹

Co-authors: Samuel Clark¹; Seth Nemesure¹; Wenge Fu

¹ *Brookhaven National Laboratory*

Corresponding Authors: fu@bnl.gov, seth@bnl.gov, sclark@bnl.gov, jmaldonad@bnl.gov

The electronic logbook (elog) system used at Brookhaven National Laboratory's Collider-Accelerator Department (C-AD) allows users to customize logbook settings, including specification of favorite logbooks. Using machine learning techniques, configurations can be further personalized to provide users with a view of entries that match their specific interests. Natural language processing (NLP) models are used to augment the elog system by classifying and finding similarities in entries. A command line interface tool is used to ease automation of NLP tasks in the controls system. A test web interface will be developed for users to enter phrases, terms, and sentences as search terms for the NLP models. The website will return useful information about a given search term. This

technique will create recommendations for each user, filtering out unnecessary results generated by current search techniques.

Primary Keyword:

AI-based controls

Secondary Keyword:

AI-based controls

Tertiary Keyword:

AI-based controls

Tools for Humans / 58

PACuna: Automated Fine-Tuning of Language Models for Particle Accelerators

Authors: Annika Eichler¹; Antonin Sulc²; Raimund Kammering³; Tim Wilksen²

¹ *Deutsches Elektronen Synchrotron DESY*

² *DESY MCS*

³ *DESY*

Corresponding Authors: raimund.kammering@desy.de, antonin.sulc@desy.de, tim.wilksen@desy.de

Navigating the landscape of particle accelerators has become increasingly challenging with recent surges in contributions. These intricate devices challenge comprehension, even within individual facilities.

To address this, we introduce PACuna, a fine-tuned language model refined through publicly available accelerator resources like conferences, pre-prints, and books.

We automated data collection and question generation to minimize expert involvement and make the data publicly available.

PACuna demonstrates proficiency in addressing accelerator questions, validated by experts.

Our approach shows adapting language models to scientific domains by fine-tuning technical texts and auto-generated corpora capturing the latest developments can further produce pre-trained models to answer some specific questions that commercially available assistants cannot and can serve as intelligent assistants for individual facilities.

Primary Keyword:

foundation models

Secondary Keyword:

Tertiary Keyword:

Methods / 26

AI/ML Coupling & Surrogates in BLAST Accelerator Modeling Codes

Authors: Axel Huebl¹; Ryan T Sandberg¹; Remi Lehe¹; Chad E Mitchell¹; Marco Garten¹; Ji Qiang¹; Jean-Luc Vay¹

¹ *Lawrence Berkeley National Laboratory*

Corresponding Authors: rlehe@lbl.gov, axelhuebl@lbl.gov

Detailed modeling of particle accelerators can benefit from parallelization on modern compute hardware such as GPUs and can often be distributed to large supercomputers. Providing production-quality implementations, the Beam, Plasma & Accelerator Simulation Toolkit (BLAST) provides multiple modern codes to cover the widely different time and length scales between conventional accelerator elements and advanced, plasma-based elements. The Exascale code WarpX provides electromagnetic and -static, t-based particle-in-cell routines, advanced algorithms and is highly scalable. For beam-dynamics, the s-based ImpactX code provides an efficient implementation for tracking relative to a nominal reference trajectory, including space charge. Yet, integrated modeling of “hybrid” beamlines integrating both detailed plasma models and large-scale transport at full detail require exchange between both codes and are limited by the computational speed of the most-detailed element, usually the plasma element.

In this work, we present an alternative approach to coupling particle-in-cell models and codes beyond direct data exchange or reduced details for accelerator modeling. In particular, we investigate and demonstrate detailed data-driven modeling based on high-quality WarpX simulations that were used to train surrogate models for the beam transport code ImpactX. We describe new workflows, illuminate predictive quality, performance and applicability to central research topics in advanced accelerator research, such as staging of laser-wakefield accelerators.

Primary Keyword:

surrogate model architecture

Secondary Keyword:

MLOps

Tertiary Keyword:

digital twins

Methods / 82

Experience with ML-based Model Calibration Methods for Accelerator Physics Simulations

Authors: Auralee Edelen¹; Christopher Mayes²; Claudio Emma²; Frederick Cropp²; Juan Pablo Gonzalez-Aguilera³; Kathryn Baker⁴; Ryan Roussel⁵; Sanjeev Chauhan⁶; Tobias Boltz⁵

¹ *SLAC, Stanford*

² *SLAC*

³ *University of Chicago*

⁴ *STFC / ISIS*

⁵ *SLAC National Accelerator Laboratory*

⁶ *Duke University*

Corresponding Authors: cemma@slac.stanford.edu, cmayes@slac.stanford.edu, ericcropp@physics.ucla.edu, sanjeev.chauhan@duke.edu, k.baker@stfc.ac.uk, tboltz@slac.stanford.edu, rroussel@slac.stanford.edu, jpga@uchicago.edu, edelen@slac.stanford.edu

Physics simulations of particle accelerators enable predictions of beam dynamics in a high degree of detail. This can include the evolution of the full 6D phase space distribution of the beam under the impact of nonlinear collective effects, such as space charge and coherent synchrotron radiation. However, despite the high fidelity with respect to the expected beam physics, it is challenging to obtain simulations that predict the behavior of the as-built system with a high degree of accuracy

due to the numerous sources of static errors and time-varying changes. Identifying and accounting for sources of mismatch between simulation and measurement, i.e. “model calibration”, is often a laborious process conducted by expert physicists, and it frequently requires time-intensive gathering of data, such as extensive parameter scans. ML-based approaches for model calibration can potentially leverage larger, more diverse sets of data and can consider a wider variety of possible sources of error simultaneously. Information-based sampling can also improve the efficiency of the data acquisition. Here we discuss the results of several model calibration approaches we have applied to common beamline setups in accelerators, as well as LCLS, LCLS-II, and FACET-II. We also review future plans for regular deployment of these methods to keep online models up-to-date as these machines change over time. The approaches center around the use of learnable calibration functions applied to differentiable models that represent the expected physics behavior, including both neural network surrogate models trained on physics simulations and directly differentiable physics simulations. Furthermore, we comment on uncertainty quantification for the learned calibration factors and disentangling multiple possible error sources. Finally, we discuss the software tools we have been working on to facilitate model calibration.

Primary Keyword:

surrogate model architecture

Secondary Keyword:

active learning

Tertiary Keyword:

digital twins

Methods / 44

ML methods for noise reduction in industrial LLRF systems

Authors: Jonathan Edelen¹; Morgan Henderson¹

Co-authors: Auralee Edelen²; Christopher Hall¹; Jorge Diaz Cruz³; Joshua Einstein-Curtis¹

¹ *RadiaSoft LLC*

² *SLAC National Accelerator Laboratory*

³ *SLAC*

Corresponding Authors: mhenderson@radiasoft.net, dejorge@slac.stanford.edu, chall@radiasoft.net, joshec@radiasoft.net, jedelen@radiasoft.net, edelen@slac.stanford.edu

Typical operational environments for industrial particle accelerators are less controlled than those of research accelerators. This leads to increased levels of noise in electronic systems, including radio frequency (RF) systems, which make control and optimization more difficult. This is compounded by the fact that industrial accelerators are mass-produced with less attention paid to performance optimization. However, growing demand for accelerator-based cancer treatments, imaging, and sterilization in medical and agricultural settings requires improved signal processing to take full advantage of available hardware and increase the margin of deployment for industrial systems. To that end, we applied several machine learning algorithms and one Bayesian filtering algorithm for removing noise from RF signals. We explored the application of different types of autoencoders including recurrent, convolutional, and variational autoencoders. In addition we developed a Kalman filter algorithm to be used as a non-ML benchmark. Our methods were first developed using simulation and then applied to measurement data from an industrial LINAC under development at RadiaBeam. This talk provides an overview of our methods and a statistical analysis of their performance on the simulation data. We will then show results on measurement data prior to and after model retraining.

Primary Keyword:

timeseries forecasting

Secondary Keyword:

AI-based controls

Tertiary Keyword:

anomaly detection

Methods / 47

Leveraging Differential Algebraic Methods for Enhanced Beam Dynamics Simulation with Machine Learning

Author: Chong Shik Park¹

¹ *Korea University, Sejong*

Corresponding Author: kuphy@korea.ac.kr

This study introduces an innovative approach that harnesses machine learning in conjunction with Differential Algebraic (DA) techniques to simulate beam dynamics in particle accelerators. Beam dynamics simulations are complex, involving high-dimensional phase spaces and intricate equations of motion. By integrating the DA method, which deals with function derivatives, with machine learning, we develop a novel framework that can efficiently model and predict beam behavior.

Our approach combines deep neural networks with DA to create a machine learning model capable of learning the underlying physics and dynamics of the accelerator system, including space charge and non-linear effects. The resulting model offers accelerated simulations and enables real-time optimization of accelerator parameters.

Several case studies showcase the effectiveness of this approach, revealing improved insight into system behavior. By bridging traditional symbolic methods with machine learning, this research propels accelerator physics into a new era. This synergy between DA and machine learning promises more accurate and efficient simulations, enhancing accelerator design, optimization, and real-time control for a wide range of scientific and industrial applications.

Primary Keyword:

differentiable models

Secondary Keyword:

Tertiary Keyword:

Methods / 49

Cheetah –A High-speed Differentiable Beam Dynamics Simulation for Machine Learning Applications

Authors: Jan Kaiser¹; Chenran Xu²; Annika Eichler³; Andrea Santamaria Garcia⁴

¹ *Deutsches Elektronen-Synchrotron DESY*

² *Karlsruhe Institut für Technologie (KIT)*

³ *DESY*

⁴ *Karlsruhe Institute of Technology*

Corresponding Authors: andrea.santamaria@kit.edu, annika.eichler@desy.de, jan.kaiser@desy.de, chenran.xu@kit.edu

Machine learning has emerged as a powerful solution to the modern challenges in accelerator physics. However, the limited availability of beam time and the high computational cost of simulation codes pose significant hurdles in generating the necessary data for training state-of-the-art machine learning models. Furthermore, optimisation methods can be used to tune accelerators and perform complex system identification tasks. However, they too require large numbers of samples of expensive-to-compute objective functions in order to achieve state-of-the-art performance. In this work, we introduce *Cheetah*, a PyTorch-based high-speed differentiable linear-beam dynamics code that enables fast collection of large datasets and sample-efficient gradient-based optimisation, while being easy to use, straightforward to extend and integrating seamlessly with widely adopted machine learning tools. Ultimately, we believe that Cheetah will simplify the development of machine learning-based methods for particle accelerators and fast-track their integration into everyday operations of accelerator facilities.

Primary Keyword:

differentiable models

Secondary Keyword:

digital twins

Tertiary Keyword:

Methods / 87

AI-assisted design of final cooling system for a muon collider

Author: Elena Fol¹

¹ CERN

Corresponding Author: elena.fol@cern.ch

The design of novel accelerator components such as ionisation cooling channel for a muon collider necessitates extensive simulation and optimization studies. We present the application of Bayesian Optimization and surrogate model - based evaluation of lattice parameters, which allowed to surpass the baseline cooling performance. Robust emittance estimation throughout the cooling channel is crucial when optimising the cooling of non-gaussian, correlated muon beams. We compare various anomaly detection algorithms applied to 6D phase space, show how these techniques can enhance the emittance estimation and discuss potential further application of other unsupervised learning methods in the design studies of future facilities.

Primary Keyword:

ML-based optimization

Secondary Keyword:

anomaly detection

Tertiary Keyword:

General / 96

4th ICFA MaLAPA Summary

Author: Daniel Ratner¹

¹ *SLAC*

Corresponding Author: dratner@slac.stanford.edu

Summary of this year's ICFA Beam Dynamics Mini-Workshop on Machine Learning Applications for Particle Accelerators.

General / 97

Closing

Corresponding Authors: ihnam@postech.ac.kr, sch0914@postech.ac.kr