

ICALEPCS 2023: 3rd Data Science and Machine Learning Workshop

Report of Contributions

Contribution ID: 1

Type: **not specified**

Workshop Introduction

Saturday 7 October 2023 08:30 (20 minutes)

Primary author: GONZALEZ BERGES, Manuel (CERN)

Presenter: GONZALEZ BERGES, Manuel (CERN)

Contribution ID: 2

Type: **not specified**

Lunch Break

Contribution ID: 3

Type: **not specified**

Wrap-up

Contribution ID: 4

Type: **not specified**

Common Problems in Early Stage Projects at the ISIS Neutron and Muon Source

Saturday 7 October 2023 13:50 (15 minutes)

At the ISIS Neutron and Muon Source, we are still relatively early on in our pursuit to integrate machine learning into the operations of the accelerator. Consultation with various teams across the accelerator has highlighted three key areas where machine learning can be leveraged most effectively, namely fault diagnosis and prediction, the use of virtual diagnostics and intelligent control of the machine. However, in the case of each of these themes we have encountered complications that may limit their development or practical use that we are keen to discuss with other facilities who may have more knowledge and experience mitigating against these issues. Some of these items to be considered include high dimensional feature selection, dealing with highly correlated outputs and how to match models trained on physics simulation with live behavior of the machine.

Primary author: BAKER, Kathryn

Presenter: BAKER, Kathryn

Session Classification: Project presentations/demos

Contribution ID: 5

Type: **not specified**

Neural Networks for Anomaly Detection in LINACs, Injectors, and Transfer Lines

Saturday 7 October 2023 12:30 (20 minutes)

Maximizing up-time of accelerators relies heavily on the ability to detect and diagnose changes in the machine. The application of machine learning for anomaly detection remains a rich area of research. RadiaSoft has been developing methods for anomaly detection in collaboration with Jefferson Lab, Brookhaven National Lab, and SLAC. Here we provide a survey of recent innovations in anomaly detection for particle accelerators and present results from our recent work. Our studies are focused on the low energy injector at CEBAF, the AGS to RHIC transfer line at BNL, and industrial accelerators for radiotherapy and imaging. We focused on the use of two neural network architectures, inverse models and variational autoencoders. This talk will provide high level context for how these methods are utilized for anomaly detection and results from our studies using both simulation and measurement data.

Primary author: EDELEN, Jon (Radiasoft)

Presenter: EDELEN, Jon (Radiasoft)

Session Classification: Project presentations/demos

Contribution ID: 6

Type: **not specified**

A Potential of Use of Language Processing in Accelerator Control Systems

Saturday 7 October 2023 12:50 (20 minutes)

Particle accelerators rely on complex control systems for their operation. As accelerators grow in scale and complexity, developing and maintaining effective control systems becomes increasingly challenging. In this presentation, we will explore the potential for applying natural language processing (NLP) techniques to improve accelerator operations by closely examining the use of textual data.

We will present our applications of NLP algorithms to logbook data from DESY and BESSY. Initial results demonstrate feasibility for using NLP to automatically parse log entries, categorize events, detect problems, and surface important information.

However, challenges remain in handling physics terminology, noisy data, and model generalization. This presentation will provide an overview of how natural language processing can be applied to accelerate logbooks in field of accelerator controls.

Primary author: SULC, Antonin (DESY)

Presenter: SULC, Antonin (DESY)

Session Classification: Project presentations/demos

Contribution ID: 7

Type: **not specified**

Addressing protein serial crystallography 36 GB/s data-rate challenge with FPGAs and GPUs

Saturday 7 October 2023 13:30 (20 minutes)

Serial crystallography [1] is a technique used at synchrotrons and X-ray free electron lasers to solve protein structures from random still diffraction images of thousands of small crystals. The technique is one of the most data intensive techniques at X-ray facilities. With novel detectors, like the 9 MPixel JUNGFRÄU [2] currently commissioned at the Paul Scherrer Institute, it is possible to acquire a continuous stream of images at 36 GB/s.

Such large data rates challenge the current way images are handled in crystallography, i.e., it is no longer possible to save to disk storage every image and every pixel irrespective of their value for the scientific question [3]. On-the-fly data analysis and compression become a key to sustainable operations of high data rate detectors. Given a very high data throughput, such analysis requires computing accelerators, like field programmable gate arrays (FPGAs) and general-purpose graphical processing units (GPUs).

In this presentation, I will talk about our practical experience from implementing data science methods for on-the-fly analysis on computing accelerators [4]. I will give a practical example of spot finding algorithms that we implemented on GPUs and FPGAs (with high-level synthesis), highlighting differences in both approaches. I will also give an outlook of our early-stage developments in image analysis with machine learning methods.

[1] T. Weinert et al. (2019). *Science*, 365, 61-65.

[2] F. Leonarski et al. (2018). *Nat. Methods*, 15, 799–804.

[3] F. Leonarski et al. (2020). *Struct. Dyn.*, 7, 014305.

[4] F. Leonarski et al. (2023). *J. Synchrotron Rad.*, 30, 227.

Primary author: LEONARSKI, Filip

Presenter: LEONARSKI, Filip

Session Classification: Project presentations/demos

Contribution ID: 8

Type: **not specified**

Use of Machine learning for Denoising Beam Profile Measurements

Saturday 7 October 2023 13:10 (20 minutes)

Several CERN accelerators are being equipped with Beam Gas Ionization (BGI) profile monitors using high resolution Timepix3 detectors resulting in very powerful and not destructive measurements [1]

The images produced by these detectors contain the signal from ionization electrons as well as noise coming from different sources (mainly beam losses) and other artifacts like noisy pixels or the RF shield.

Several approaches are being studied to remove the noise and the artifacts from the images to improve the beam measurement. The presentation will give an overview of these approaches.

[1] <https://bgi.web.cern.ch/introduction/>

Primary authors: MARTINEZ SAMBLAS, Javier; GONZALEZ BERGES, Manuel (CERN)

Presenter: MARTINEZ SAMBLAS, Javier

Session Classification: Project presentations/demos

Contribution ID: 9

Type: **not specified**

Tutorial I: Linear and Logistic Regression

Saturday 7 October 2023 08:50 (40 minutes)

Primary author: VALENTINO, Gianluca (University of Malta (MT))

Presenter: VALENTINO, Gianluca (University of Malta (MT))

Contribution ID: **10**

Type: **not specified**

Coffee Break

Contribution ID: 11

Type: **not specified**

Tutorial II: Neural Networks, Unsupervised Learning and Advanced Topics

Saturday 7 October 2023 10:00 (1h 30m)

Primary author: VALENTINO, Gianluca (University of Malta (MT))

Presenter: VALENTINO, Gianluca (University of Malta (MT))

Contribution ID: 12

Type: **not specified**

Discussion Sesssion (only if enough interest)

Saturday 7 October 2023 15:00 (2 hours)

Presenter: ALL PARTICIPANTS

Contribution ID: 13

Type: **not specified**

Additional presentations and/or discussions

Saturday 7 October 2023 14:05 (25 minutes)

Session Classification: Project presentations/demos