

Workshop Title: EuroCC – Accelerated Machine Learning with Intel AI Software Tools

Date: October 25, 2022, 9am – 5:20pm CEST

Where: Online, via Webex. Details will be provided after registration.

To participate please register at <https://bsc3.typeform.com/Intelworkshop>.

Purpose of this Workshop

The Intel AI Analytics Toolkit gives data scientists, AI developers, and researchers familiar Python* tools and frameworks to accelerate end-to-end data science and analytics pipelines on Intel® architecture. The components are built using oneAPI libraries for low-level compute optimizations. This toolkit maximizes performance from pre-processing through machine learning and provides interoperability for efficient model development.

Using this toolkit, you can:

- Deliver high-performance, deep learning training on Intel® XPU's and integrate fast inference into your AI development workflow with Intel®-optimized, deep learning frameworks for TensorFlow* and PyTorch*, pretrained models, and low-precision tools.
- Achieve drop-in acceleration for data pre-processing and machine learning workflows with compute-intensive Python* packages, Modin*, scikit-learn*, and XGBoost, optimized for Intel.
- Gain direct access to analytics and AI optimizations from Intel to ensure that your software works together seamlessly.

We welcome you to join us in learning about the Intel AI Analytics toolkit and how to integrate the various optimized components into your workflow. The presenters will show you how to use the tools for both classical machine learning and deep learning workflows. We will also cover adjacent topics around this including:

- Accelerating your model development by enhancing your experimentation with SigOpt
- Federated learning using the flexible and extensible OpenFL framework
- AI-driven multiphysics HPC applications on Intel architecture. This is to address the convergence of AI and HPC applications
- Easily speed up Deep Learning inference by writing your code once and deploying it on any supported Intel hardware using the OpenVINO toolkit.

After the event the attendees should understand:

- How to accelerate end-to-end AI and Data Science pipelines, achieve drop-in acceleration with optimized Python tools built using oneAPI libraries.
- How to achieve high-performance deep learning training and inference with Intel-optimized TensorFlow and PyTorch versions, and low-precision optimization with support for fp16, int8 and bfloat16.

- Seamlessly scale Pandas workflows across multi-node dataframes with Intel® Distribution of Modin.
- Increase machine learning model throughput with algorithms in Scikit-learn and XGBoost optimized for Intel architectures.

Prerequisites

It is assumed that participants have a basic knowledge about machine learning and deep learning and should be able to program in Python.

Attendees & Fee

- The online workshop is free of charge and it is open to everyone within EuroCC and CASTIEL.
- **To participate please register at <https://bsc3.typeform.com/Intelworkshop>.**
- Amount of seats is limited.

Agenda: EuroCC – Accelerated Machine Learning with Intel AI Software Tools

Accelerated Machine Learning with Intel

9:00AM – 9:10AM: *Welcome: Present the agenda and the speakers – Edmund Preiss, Dr. Séverine Habert*

1. 9:10AM – 9:45AM: **Introduction to AI Domains and Overview of Intel Hardware and Software Acceleration for AI**, Dr. Séverine Habert

In this session, we will give an overview of AI domains and introduce the hardware features that are powering AI on Intel. You will then get a first glance at the software stack harnessing them, namely the Intel® oneAPI Analytics Toolkit.

2. 9:45AM - 10:30AM: **How to Accelerate Classical Machine Learning on Intel Architecture**, Roy Allela

In this session, we will cover the Intel-optimized libraries for Machine Learning. Python is currently ranked as the most popular programming language and is widely used in Data Science and Machine Learning. We will begin by covering the Intel® Distribution for Python and its optimizations. We will then cover the optimizations for ML Python packages such as Modin, Intel® Extension for Scikit-learn and XGBoost. The presentations will be accompanied with demos to showcase the performance speedup.

10:30AM – 10:45AM Break (15 min)

3. 10:45AM - 11:30AM: **Enhance Your Experimentation with SigOpt**, Tobias Andreason

Modeling is a scientific process that requires experimentation to get right. But experimentation is only as effective as the tools applied to it. SigOpt is an Intelligent Experimentation platform that empowers AI modelers to design experiments by asking the right questions, explore experiments to understand their modeling problems, and optimize their experiments to get the best results.

4. 11:30AM-12:15PM: **Federated Learning**, Walter Riviera

When it comes to AI, we can't really address the topic without talking about Data. Nowadays, the demand for data has increased the needs to collect fresh information to feed bigger and more ambitious AI models. However, while data is being collected at different levels, accessing it might not always be possible due to physical constraints (i.e. remote locations) or regulations in place (i.e. GDPR, HIPAA, POPIA). In this session we'll learn what is Federated Learning and how we can build a real federation that is able to leverage distributed dataset to train a shared model and solve the data access problem.

12:15PM – 1:30PM Lunch break (1h15)

Accelerated Deep Learning with Intel

5. 1:30PM - 2:15PM: **Optimize Deep Learning on Intel – Same code just faster!**, Akash Dhamasia

In this session, we present to you what is behind the scenes of Deep Learning with the highly-optimized Intel® oneDNN library in order to get the best-in-class performance on Intel hardware. We then show you Intel® oneDNN in action in DL frameworks such as the Intel-optimized TensorFlow, Intel-optimized PyTorch and the Intel® Extension for PyTorch (IPEX).

6. 2:15PM - 2:45PM: **AI-driven Multiphysics HPC Applications on Intel Architecture**, Dr. Massoud Rezavand

A major challenge in HPC is to make use of and understand the massive amounts of data that are being produced when running numerical simulations. For ML on the other hand, the challenge is to have access to enough data so that we have the confidence that our models truly understand the world. Therefore, researchers are looking to replace components of HPC applications with ML models to (a) reduce the need for data storage, (b) accelerate the simulations by ML models to capture longer timescales, and (c) achieve accurate simulations in some problems that the classical solvers are not applicable to. In this session we present this interdisciplinary field and highlight recent achievements on Intel® architectures.

2:45PM – 3:00PM Break (15min)

7. 3:00PM - 3:45PM: **Introduction to Neural Network Compression Techniques**, Dr. Nikolai Solmsdorf

In this session, we will explain various network compression techniques in Deep Learning—such as quantization, pruning, and knowledge distillation—, their benefits in terms of performance speed-up, and finally we will showcase you the Intel tools that help you compress your model, like the Intel® Neural Compressor.

8. **3:45PM - 4:30PM: Easily speed up Deep Learning inference – Write once deploy anywhere!**, Vladimir Kilyazov

In this session, we will showcase the Intel® Distribution of OpenVINO™ Toolkit that allows you to optimize for high-performance inference models that you trained with TensorFlow* or with PyTorch*. We will demonstrate how to use it to write once and deploy on multiple Intel hardware.

4:30PM – 5:00PM Q&A, Closure (30min)

- To participate please register at <https://bsc3.typeform.com/Intelworkshop>.