The ML_INFN Initiative

Luca Giommi
Istituto Nazionale di Fisica Nucleare
CNAF
The ML_INFN Initiative

INFN Research and structures

216 activities distributed in 33 structures (labs, groups and divisions)

Particle Physics
- 17 experiments

Astroparticle Physics
- 45 experiments

Nuclear Physics
- 23 experiments

Theoretical Physics
- 35 initiatives

Technological Research
- 96 experiments
Machine Learning applications in INFN

Most of the experiments and initiatives produce, analyse or process digital data.

Enthusiasm on the modern data processing technologies!

Gravitational wave detection

Raw radiation detector data

Theoretical computations on the lattice

Research on innovative imaging technologies
ML_INFN: The structure of the project

Applications of Machine Learning
HEP, MedPhys, GW detection, Theory...

Infrastructure
- NVIDIA CUDA
- Jupyter
- Docker

Stewardship
- K
- TensorFlow
- Scikit learn

Knowledge Base
- INFN Cloud
- MinIO
- IAM

Virtualization and orchestration layer
developed and maintained by INFN Cloud
WP1. The infrastructure
INFN Cloud

ML_INFN is built on top of INFN Cloud: a data-lake centric, heterogeneous federated Cloud infrastructure spanning over multiple sites across Italy, providing an extensible portfolio of solutions tailored to multidisciplinary scientific communities.

See also...

Talks: Fanzago, Marcon, Mazzitelli

Poster: Michelotto, Sinisi
Federated bare-metal resources

1× SuperMicro + 1× E4 servers:

- 1.8 TB RAM
- 64-128 CPU cores
- 36 TB local storage (NVMe)
- 8× Tesla T4 GPUs
- 5× RTX 5000 GPUs
- 1× A30 GPU
- 1× A100 GPU, served as 7 independent MIG slices
- 10 GbE connection to CNAF resources

Federated to CNAF OpenStack and INFN Cloud

Storage solutions

CERN experiments data, contained in INFN Tier-1 storage, are remotely accessed via NFS

Hypervisors integrate Ceph to manage persistent virtual volumes accessed from the VM via POSIX
WP2. Stewardship
The ML_INFN Initiative

Machine Learning hackathons: Base and Advanced level

To foster the adoption of machine learning tools and techniques in INFN community, we organize events to discuss ML algorithm with the time to look at (and hack) the code.

Starting-level Hackathons
Jun 2021, Dec 2021, Jun 2023
- online events with no fee
- up to 60 participants
- 1 tutor per 5 participants
- INFN Cloud CPUs with shared filesystem

Advanced Hackathons
Nov 2022
- in-person events
- up to 30 participants
- (almost) 1 tutor per participant
- INFN Cloud GPUs with shared filesystem
Base hackathon: Lecture Program

Day 1

- Theoretical introduction to ML
- Cloud and Cloud Resources

Day 2

- Neural Networks
- Deep Neural Network Applications to INFN research

Day 3

- Exercises with tutors continuous support

Lunch break

- Hands-on
  - Numpy, Pandas and Keras
  - Exercises with tutors on demand

Closure

- Reports from the students
Hackathon use cases: 10 groups, one tutor per group

Jet b-tagging at CMS
Recurrent Neural Networks with LSTM

Higgs searches at CMS
Deep Neural Networks and Advanced Keras

Gravitational Waves with Virgo
Autoencoders, anomaly detection and compression

Segmentation of CT scans
Convolutional Neural Networks Handling 2D and 3D datasets
**Advanced hackathon: Lecture Program**

**Day 1**
- Lectures: Advanced Models (U-Nets, GANs, NFs, ...)
- Hands-on: Beyond Keras (Coding lower-level ops)

**Day 2**
- Lectures: Advanced Models (Transformers, GNNs)
- Hands-on: Explainability

**Day 3**
- Lectures: Ongoing R&Ds in Machine Learning (FPGAs, HPO, ...)
- Hands-on: BondMachine Compiling NN in VHDL

**Day 4**
- Lectures: Continuous support

---

**Lunch break**

**Lectures**
- Cloud Infrastructure and High Performance Computing

**Hackathon**
- Exercises with tutors continuous support

---

Luca Giommi (INFN CNAF, Italy)
Advanced hackathon use cases

**Lung Segmentation with U-Nets**

- U-Nets, custom loss, custom data loaders

**Domain Adaptation in HEP**

- Adversarial Training, Gradient Tape

**GNN and Transformers in HEP**

Application to Jet Tagging

**Explainability**

- Shapley and GradCAM
WP3. The Knowledge Base
Atlassian Confluence was used to build a Knowledge Base reporting several machine-learning use cases, including those discussed at the hackathon.

Each entry includes:

- **Runnable example** as a Jupyter notebook or a git repository
- **Contact information** of one or more experts

<table>
<thead>
<tr>
<th>Name and Link</th>
<th>ML Technologies</th>
<th>Scientific Field</th>
<th>ML Tools</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>B tagging in CMS (template version)</td>
<td>CNN, LSTM</td>
<td>High Energy Physics</td>
<td>Keras + Tensorflow</td>
<td>Realistic application</td>
</tr>
<tr>
<td>LHCb Masterclass with Keras</td>
<td>DE, MLP</td>
<td>High Energy Physics</td>
<td>ROOT + Keras + TF</td>
<td>Introductory tutorial</td>
</tr>
<tr>
<td>MNIST in C header</td>
<td>MLP</td>
<td></td>
<td>Keras</td>
<td>Free-styling tutorial</td>
</tr>
<tr>
<td>LUMIN: Lumin Unifies Many Improvements for Networks</td>
<td>CNN, RNN, GNN</td>
<td>High Energy Physics</td>
<td>PyTorch</td>
<td>Package use examples</td>
</tr>
<tr>
<td>INFERN0: Inference-Aware Neutral Optimisation</td>
<td>NN</td>
<td>High Energy Physics</td>
<td>Keras + Tensorflow</td>
<td></td>
</tr>
<tr>
<td>An introduction to classification with PyTorch</td>
<td>Fisher, BOT, MLP</td>
<td>High Energy Physics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This section of the ML-INFN Confluence Space contains the Knowledge Base of fully implemented use cases. It has been created in order to provide new users getting close to machine learning with concrete examples, with step by step guides for reproducibility. The division into categories is multidimensional:

- Dimension 1: per machine learning technology (CNN, Auto encoders, LSTM, GraphNet, ...)
- Dimension 2: per scientific field (High Energy Physics, Gravitational Waves, Medical Physics, ...)
- Dimension 3: per type of used tool

and is implemented via Confluence labels.
Seminars organized in 2023

1) Improving parametric neural networks for high-energy physics (and beyond)
2) Cell counting with cell-ResUnet
3) A Neural-Network-defined Gaussian Mixture Model for particle identification applied to the LHCb fixed-target programme
4) Deep-learning emulators and hierarchical Bayesian inference: application to gravitational-wave astronomy
5) New Physics Learning Machine (NPLM): a tool for statistical anomalies detection in presence of systematic uncertainties
6) Machine Learning as a Service for High Energy Physics (MLaaS4HEP): a service for ML-based data analyses
7) Ante-hoc explainability methods: the ProtoPNet architecture and its application on DBT images
<table>
<thead>
<tr>
<th>Title</th>
<th>Journal/Conference</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring Analytic Gradients of General Quantum Evolution with the</td>
<td>QUANTUM-AUSTRIA</td>
<td>2021</td>
</tr>
<tr>
<td>Stochastic Parameter Shift Rule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The novel Mechanical Ventilator Milano for the COVID-19 pandemic</td>
<td>PHYS FLUIDS</td>
<td>2021</td>
</tr>
<tr>
<td>How to enhance quantum generative adversarial learning of noisy</td>
<td>NEW J PHYS</td>
<td>2021</td>
</tr>
<tr>
<td>information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep learning method for TomoTherapy Hi-Art: prediction three-</td>
<td>RADIOTHER ONCOL</td>
<td>2021</td>
</tr>
<tr>
<td>dimensional dose distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model compression and simplification pipelines for fast deep neural</td>
<td>EUR PHYS J C</td>
<td>2021</td>
</tr>
<tr>
<td>network inference in FPGAs in HEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalization in Quantum Machine Learning: A Quantum Information</td>
<td>PRX QUANTUM</td>
<td>2021</td>
</tr>
<tr>
<td>Standpoint</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model compression and simplification pipelines for fast deep neural</td>
<td>EUR PHYS J C</td>
<td>2021</td>
</tr>
<tr>
<td>network inference in FPGAs in HEP (vol 81, 969, 2021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calorimetric Measurement of Multi-TeV Muons via Deep Regression</td>
<td>EUR PHYS J C</td>
<td>2022</td>
</tr>
<tr>
<td>A Neural-Network-defined Gaussian Mixture Model for particle</td>
<td>J INSTRUM</td>
<td>2022</td>
</tr>
<tr>
<td>identification applied to the LHCb fixed-target programme</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tau Lepton Identification With Graph Neural Networks at Future</td>
<td>FRONT PHYS-LAUSANNE</td>
<td>2022</td>
</tr>
<tr>
<td>Electron-Positron Colliders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applications of artificial intelligence in stereotactic body</td>
<td>PHYS MED BIOL</td>
<td>2022</td>
</tr>
<tr>
<td>radiation therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model independent measurements of standard model cross sections</td>
<td>EUR PHYS J C</td>
<td>2022</td>
</tr>
<tr>
<td>with domain adaptation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust quantum classifiers via NISQ adversarial learning</td>
<td>NAT COMPUT SCI</td>
<td>2022</td>
</tr>
<tr>
<td>Toward artificial-intelligence assisted design of experiments</td>
<td>NUCL INSTRUMENT METH A</td>
<td>2023</td>
</tr>
<tr>
<td>Hyperparameter Optimisation of Artificial Intelligence for Digital</td>
<td>LECT NOTES COMPUT SC</td>
<td>2022</td>
</tr>
<tr>
<td>RESToration of Cultural Heritages (AIRESC-H) Models</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

The ML_INFN initiative has been providing many INFN experiments with the hardware and the knowledge base to assess the potential benefit of machine learning to their research for three years.

The ML_INFN project relies on INFN Cloud solutions and it federates resources optimized for ML performance in interactive and batch-like usage patterns (high-end professional GPUs, NVMe disks, many-core high-RAM systems)

A series of national training events (machine learning hackathons) and a collection of tutorials and real applications within the INFN community (knowledge base) contribute to building a network of experienced and enthusiast machine learning practitioners, lowering the skill gap to benefit from machine learning developments.
Thanks for the attention
Questions?
Backup
The numbers of ML_INFN

12 INFN structures involved in the developments, training activities and hackathons

79 researchers devoting a fraction of their time to promote ML techniques for research

14 professional GPUs made available and accessible through the INFN Cloud Interface

143 participants to the hackathons, ranging from students to permanent staff members
Stories of success [1]: building template models for LHCb

ML_INFN infrastructure was used to develop a model for the Particle Identification response of the LHCb detector as a Gaussian-Mixture model.

With Gaussian parameters inferred with a Deep Neural Network.

Traditional method based on reweighted MC

Deep Learning model

S. Mariani et al, “A Neural-Network-defined Gaussian Mixture Model for particle identification applied to the LHCb fixed-target programme”, JINST 17 (2022) P02018
Stories of success [2]: studying LQCD with CNNs

A Deep Neural Network is trained in a semi-supervised manner to define an effective order-parameter for Gauge theory where a real order-parameter is not defined.

The study was made possible thanks to the GPUs provided by the ML_INFN initiative.

A. Palermo, M.P. Lombardo et al. “Machine learning approaches to the QCD transition”, Proceeding of LATTICE21
Stories of success [3]: X-rays to visible colors for CH

X-ray fluorescence spectroscopy widely used for Heritage Conservation and non-invasive probe of pictorial artworks.

Deep Neural Network models are trained to reconstruct the image from the XRF scan of the pixels.

Public repository Hackathon tutorial