Spoke 2 - Fundamental Research & Space Economy

Proposal for WP2

September 22, 2022

N. De Filippis on behalf of the team at POLIBA

WP2 goals and timeline

2) Experimental High Energy Physics: selection, data reduction, simulation and reconstruction algorithms (either via explicit programming or large-scale Machine Learning solutions) for HEP experiments (LHC, Future Colliders, KEK, IHEP, neutrino experiments...), with applications ranging from innovative triggers to distributed analysis techniques.

We plan to carry out the 3 years' research following a standard approach, in which 4 phases have been identified. For the "Scientific Work Packages" 2.1, 2.2, and 2.3 this means:

1. planning and identification: the first year of the project is dedicated to a landscape recognition for interesting state-of-the-art use case; its outcome is a work plan identifying the activities on which the core part of the project will be focusing - in particular, algorithms and services to be accomplished;
2. a realization phase, in which the actual development is performed via the staff/ hired personnel; the outcome is usable algorithms / services, documented (alpha/beta level) and ready to be tested on a larger scale;
3. a validation phase, in which the outcomes of the realization phase are verified in testbeds and proofs of concept, and benchmarked in order to assess their adherence to the specifications;
4. a wrap-up phase, in which results are analysed and consolidated in reports and white papers to be used as guidelines for similar use cases.

WP2 team at Poliba

Contact: N. De Filippis (PA)

Members:
N. De Filippis (PA) – CMS and FCC experiments
G. Bruno (PA) – ALICE experiment
G. Pugliese (PA) – CMS experiment
D. Creanza (PA) – CMS experiment
D. Colella (RTDA) – ALICE experiment

Proposals:

ML for Particle identification @ FCC-ee

Contatti: Nicola De Filippis (Poliba), Angela Taliercio (Northw. U.), E. Gorini (UniSalento)

Descrizione: In addition to tracking and calorimetry, Particle IDentification (PID) is a crucial aspect of most particle physics experiments. PID strategies and methods used by the large LHC experiments provide outstanding examples of the state-of-the-art. Along the path defined by the European Strategy for Particle Physics an electron-positron Higgs factor is the highest priority next collider. The FCC program at CERN combine in the same 100km infrastructure a high luminosity Higgs and Electroweak factory e+e- collider, followed by a 100 TeV hadron collider. The IDEA project, as proposal for an experiment along then electron-positron collider, includes an ultralight drift chamber as the main tracking device designed to provide efficient tracking, high precision momentum measurement and excellent particle identification. The charged particle identification for pions, kaons, protons, muons and electrons, based on the sub-detectors response, is considered as a machine learning problem solved in different modes: one-vs-rest, one-vs-one and multi-classification, which affect the models training and prediction. Deep neural networks and graph neural networks could be the best tools to combine the information from various detectors effectively. Moreover, the complexity of the detector and richness of the detection techniques make PID an interesting area of research also for the computer science community. The performance of the machine learning techniques for PID are also measured in terms of signal to background discrimination in the context of physics analysis for heavy flavor measurements.

Esperimento: FCC

Tecnologie: Centro calcolo ReCaS con CPU, GPU

ML for 4D Vertexing @ HL-LHC

POLIBA interested to ML for tracking

Contatti: Vieri CANDELISE (CMS), Giuseppe DELLA RICCA (CMS), Valentina ZACCOLO (ALICE)

Descrizione: The inclusion of the timing information in the event reconstruction of present and future particle detectors in High Energy Physics is expected to provide great advantages in sensitivity for many precision physics measurements and searches for new particles, when facing very high rates of particle collisions at high luminosity accelerators. The integration of a timing detector in modern experiments is expected to have a huge impact on the heavy-flavour tagging of beauty and charm quarks. For Deep Neural Networks based tagging algorithms, the inclusion of timing information as input variable for the DNN can translate into a tagging efficiency gain of up to 4-6% at constant background rejection of light flavours of 1% from the spatial separation of spurious secondary vertices coming from out of time pile-up contamination of neutral particles in the event. Based on the physics case of interest, this gain can be even stronger when multi-jet final states are the dominant process, i.e. di-Higgs hadronic decays, search for long-lived particles and new resonances decaying hadronically. Timing information will also be relevant for studying multi-charm hadrons and exotic states in high-multiplicity and high-density conditions, such as those produced in heavy-ion collisions, and to separate charge deposits to search for new (anti-)nuclei, hyper-nuclei and super-nuclei.

Esperimento: ALICE, CMS

Coordination/collaboration with UNIBA and INFN Bari foreseen

Proposals:

Big Data for High Energy Physics: use cases from the data analysis chain for the CMS experiment

Contatti: Nicola De Filippis (Poliba)

Descrizione:

The experimental high energy physics was pioneering the development of technologies to handle data for many decades in the recent past. By contrast currently big companies take the leadership in deployment, deployment and large scale production of data related solutions and the scientific community is making an effort to cope with the new tools, the software and development. Hundreds of physicists store, archive, analyze an increasing large data sample with different goals at the same time. The reseracher is expected to face the challenges imposed today and in the future for data analysis and emphasize the use of “Big data” tools for data manipulation and reception, machine learning techniques and more. New directions investigated for the different purposes will be studied and comparisons between the results by using current and new tools will be performed.

Stato: Avviato

Esperimento: CMS

Tecnologie: ReCaS, JupyterLab, Apache Spark, Analytics, GPU

Stato di Avanzamento: 1

Personale: RTDA

Coordination/collaboration with UNIBA and INFN Bari foreseen