SMARTHEP

REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

# ESR2: Real-time Rule Induction in Fraud Detection and HEP


## Laura Boggia

Supervised by Christian de Sainte-Marie, Bogdan Malaescu, Anja Butter



SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

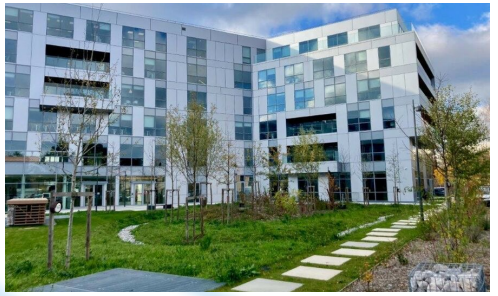

MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

# My Background

- Laura Boggia
- 06/09/1998
- Swiss & Italian
- 2017-2020: BSc in Physics at EPFL
- 2020-2022: MSc in Physics at ETH
  - Focus on theoretical physics, e.g. QFT and GR
  - Thesis on Quantum ML for HEP with IBM Research Zurich
- 2022: PhD with IBM France & Sorbonne Université
  - Supervised by Christian de Sainte-Marie & Bogdan Malaescu
  - Started 03/10/2022



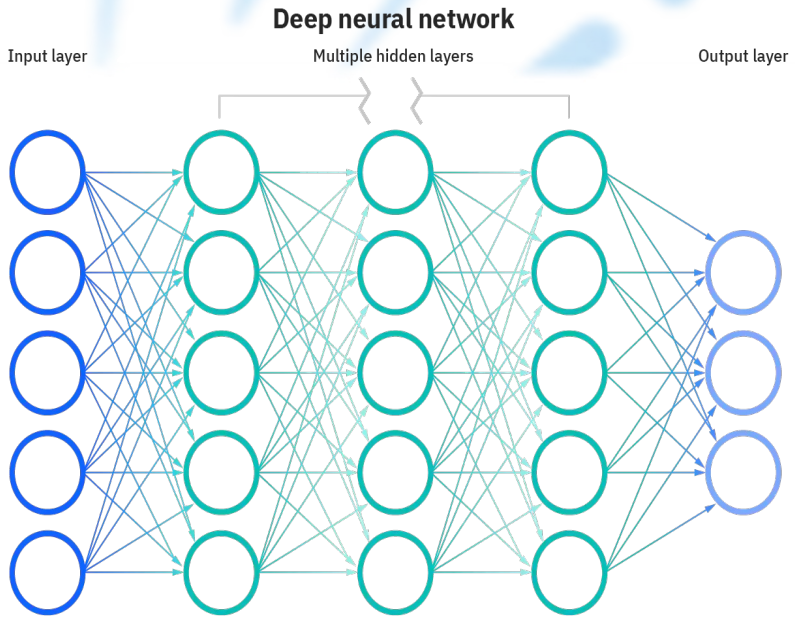



SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

# Explainable Machine Learning

- Despite success of Neural Networks (NN) their approach raises interpretability and explainability challenges
  - Very hard to understand how/why they reach a conclusion
  - For critical applications you cannot blindly trust NN models (e.g. fraud detection)
- Various approaches to make ML models more interpretable
  - Combine statistical with symbolic models, e.g. rule induction
  - Inject knowledge in statistical models, e.g. PINN




SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

# *Anomaly Detection with Jets for ATLAS*

- HEP needs new/more efficient approaches for data analysis
  - Bunch crossing every 25 ns at LHC, each bunch contains ~$10^{11}$ protons[1]
  - 3 200 TB raw data / year at ATLAS[1]
- Goal is to combine ML with anomaly detection for HEP using knowledge-based models and potentially use these models for fraud detection
- Oct-Nov 2022: Bibliographic study on jet reconstruction and calibration, online lectures on jet physics and introduction to ML
- Nov-Dec 2022: Reading up on interpretable ML and rule learning, online lectures on deep learning and scientific integrity
- Jan 2023: Starting qualification task for ATLAS
  - Improving jet calibration with ML

[1] *The ATLAS Experiment (2011): 'ATLAS fact sheet'. CERN*

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

# Training & Outreach during PhD

- Oct 2022: ‘Fête de la Science’
  - Participation at ‘my thesis in 5 minutes’
  - Guided tours of the lab for the public

- Nov 2022: ‘ATLAS Induction Day and Software Tutorial’
  - Introduction to software necessary for data analysis
  - Step-by-step guided analysis of one physics example

SMARTHEP is funded by the European Union’s Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

# Career Expectations

- Still unsure whether to pursue a career in the private industry or academia
- Goal of PhD is to gain more experience in the field of academic research

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

# *Thank you for your attention!*

Laura Boggia
ESR 2
laura.boggia@cern.ch

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086