# Data Science & Machine Learning Workshop

## Introduction

Manuel GONZALEZ-BERGES (CERN)
Gianluca VALENTINO (University of Malta)

7th October 2023

# *Some history*

# ICALEPCS 2019 – New York

Data Science and Machine Learning Workshop

-

Program Introduction

M. Gonzalez-Berges, M. Lonza

Data Science and Machine Learning Workshop - October 6, 2019

ICALEPCS2019 - New York, October 5-11, 2019

# ICALEPCS 2021 - Shangai

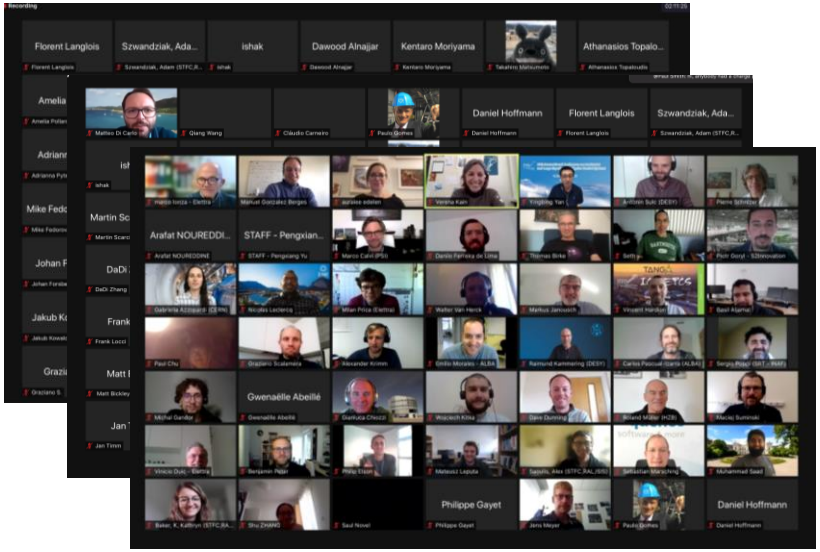

Introduction

Data Science and Machine Learning Workshop

@ICALEPCS2021

Manuel Gonzalez-Berges (CERN)
Marco Lonza (Elettra)






***130 attendees***
***( >200 registered)***

# Tracks with ML related papers

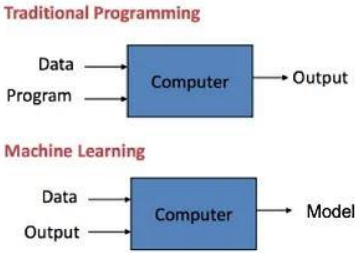
| Conferences | Tracks |
|---|---|
| ICALEPCS 2013 | • Knowledge-based Techniques |
| ICALEPCS 2015 | • Feedback Systems, Tuning |
| ICALEPCS 2017<br>(first time ML mentioned in descriptions) | • Data Analytics<br>• Feedback Control and Process Tuning |
| ICALEPCS 2019 | • Data Analytics<br>• Feedback Control and Process Tuning<br>• Experiment Control |
| ICALEPCS 2021 | • Data Analytics<br>• Feedback Control, Machine Tuning & Optimization<br>• Experiment Control<br>• Timing Systems, Synchronization & Real-Time Apps |
| ICALEPCS 2023 | Many tracks (~35 papers mention ML)<br>Specially:<br>• Artificial Intelligence & Machine Learning |

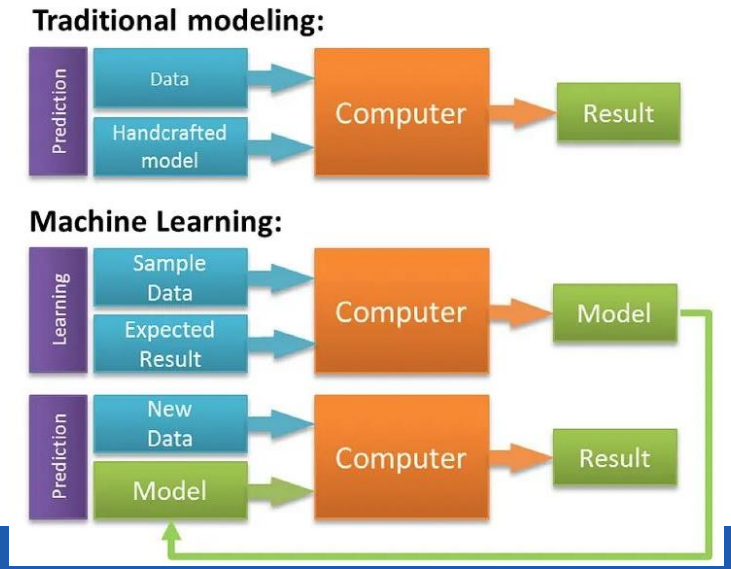
# *Some concepts*

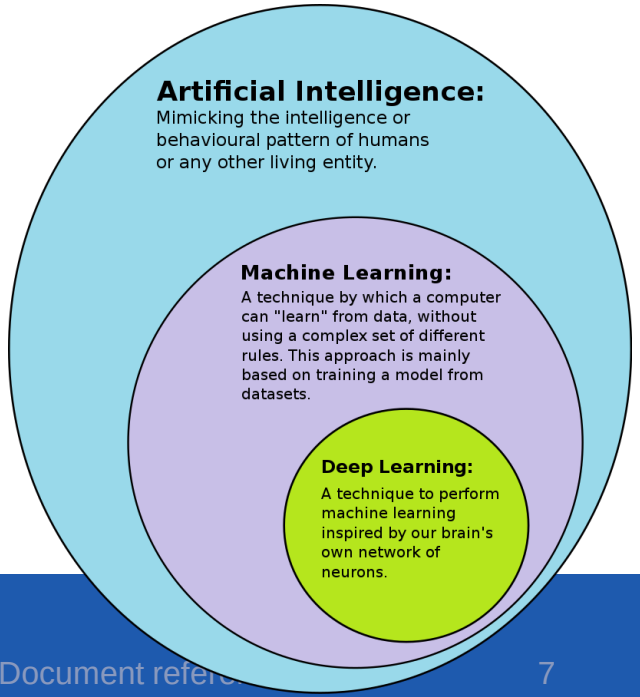
# What is Machine Learning?

WIKIPEDIA
The Free Encyclopedia

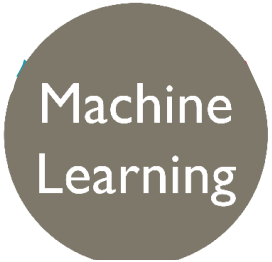
**Machine learning** (**ML**) is a field devoted to understanding and building methods that let machines "learn" – that is, methods that leverage data to improve computer performance on some set of tasks.[1]



# ML types




https://www.linkedin.com/pulse/business-intelligence-its-relationship-big-data-geekstyle/

# Workshop Program

# Participants

- Conference registration ~60
  - 41 Indico registrations (13 in the last 2 weeks)

| DS/ML Experience | |
|---|---|
| New field to me | 17 |
| Some experience | 9 |
| Regular experience | 15 |

| Programming level | |
|---|---|
| No knowledge | 1 |
| Beginner programmer | 7 |
| Regular programmer | 22 |
| Expert programmer | 11 |

# Participants Interests

- Anomaly detection
- NLP
- Computer Vision
- Denoising
- Predictive Maintenance
- Parameter optimization
- Advanced control
- Calibration
- Time series modelling
- Real-world applications
- etc

# Morning

08:00

**Welcome Coffee**

*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 08:00 - 08:30

**Workshop Introduction** *Manuel Gonzalez Berges*

*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 08:30 - 08:50

**Tutorial I: Linear and Logistic Regression** *Gianluca Valentino*

09:00

*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 08:50 - 09:30

**Coffee Break**

*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 09:30 - 10:00

10:00

**Tutorial II: Neural Networks, Unsupervised Learning and Advanced Topics** *Gianluca Valentino*

11:00

*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 10:00 - 11:30

**Lunch Break**

12:00

*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 11:30 - 12:30

# Afternoon

13:00

**Neural Networks for Anomaly Detection in LINACs, Injectors, and Transfer Lines** — *Jon Edelen*
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 12:30 - 12:50

**A potential of use of Language Processing in Accelerator Control Systems** — *Antonin Sulc*
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 12:50 - 13:10

**Use of Machine learning for Denoising Beam Profile Measurements** — *Javier Martinez Samblas*
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 13:10 - 13:30

**Addressing protein serial crystallography 36 GB/s data-rate challenge with FPGAs and GPUs** — *Filip Leonarski*
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 13:30 - 13:50

**Common Problems in Early Stage Projects at the ISIS Neutron and Muon Source** — *Kathryn Baker*
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 13:50 - 14:05

14:00

**Additional presentations and/or discussions**
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 14:05 - 14:30

**Coffee Break**
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 14:30 - 15:00

15:00

**Discussion Sesssion (only if enough interest)** — *All participants*
*CENTURY CITY CONVENTION CENTRE4 Energy Lane, Century City, Cape Town* 15:00 - 17:00

16:00

17:00

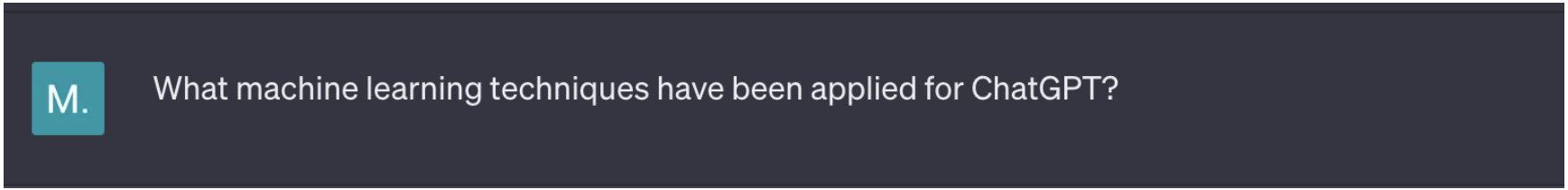
# Tutorial Sessions

Gianluca Valentino

*University of Malta*

Prof. Gianluca Valentino is an associate professor at the Department of Communications and Computer Engineering at the University of Malta, where he teaches in machine learning and pattern recognition.
He is involved in several research projects which involve the application of these techniques in various domains, from particle accelerators to earth observation, aerospace and financial data.
He spent six years with the Beams department at CERN, first as a PhD student working to automate the collimator beam-based alignment procedure, and then as a postdoctoral fellow.
He is currently a Visiting Scientist at CERN.

# Exercise


M.
What machine learning techniques have been applied for ChatGPT?


**International Organizing Committee**
Daniel Ratner (Chair) (SLAC)
Andreas Adelmann (PSI)
Ilya Agapov (DESY)
Kevin Brown (BNL)
Paul Chu (NJU)
Nobuhisa Fukunishi (RIKEN)
Kevin Li (CERN)
Hirokazu Maesaka (RIKEN)
Tia Miceli (Fermilab)

**Scientific Organizing Committee**
Tia Miceli (Chair) (Fermilab)
Andreas Adelmann (PSI)
Myung-Hoon Cho (PAL)
Paul Chu (NJU)
Auralee Edelen (SLAC)
Annika Eichler (DESY)
Yuan Gao (BNL)
Andrea Santamaria Garcia (KIT)
Georg Hoffstaetter (Cornell U.)
Yi Jiao (CAS / IHEP)
Verena Kain (CERN)
Yung-Kyun Noh (Hanyang U.)
Malachi Schram (TJNAL)
Jason St. John (Fermilab)
Tetsuhiko Yorita (Osaka U. / RCNP)

**Local Organizing Committee**
Inhyuk Nam (Chair) (PAL)
Myung-Hoon Cho (PAL)
Abin Hwang (PAL)
Gyujin Kim (PAL)
Nayoung Kim (PAL)
Chi Hyun Shim (PAL)
Haeryong Yang (PAL)
Gunsu Yun (POSTECH)

**Topics**
1. Analysis & Diagnostics
2. Anomaly Detection / Failure Prediction
3. Infrastructure / Deployment Workflows
4. Optimization & Control
5. Modeling Approaches
6. Lessons Learned

**Tutorials**
1. Reinforcement Learning
2. Model Adaptation / Up-keep
3. Transformers for Timeseries Prediction

**Call for Abstracts**
Presentation (20 mins) or Poster
Submission Deadline: November 1, 2023
Notification of Acceptance: December 1, 2023

PAL POHANG ACCELERATOR LABORATORY POSTECH POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

사진제공(권미정) - 경주시 관광자원 영상이미지

# Questions?