Machine Learning in the Vacuum Group

TE-TM

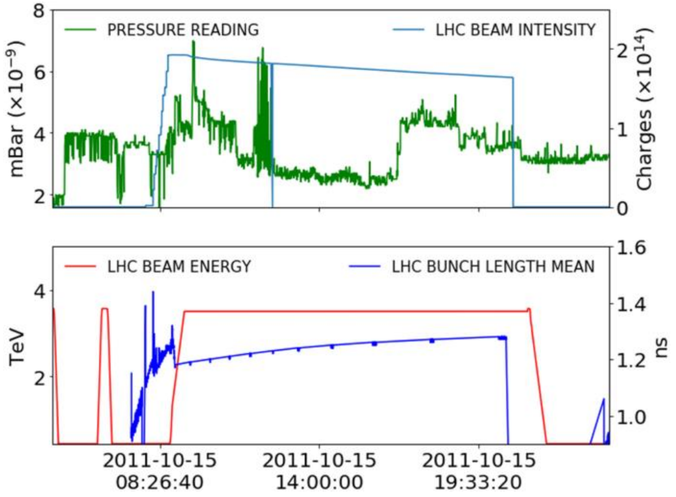
André Rocha (andre.rocha@cern.ch)

TE-VSC
Technology Department

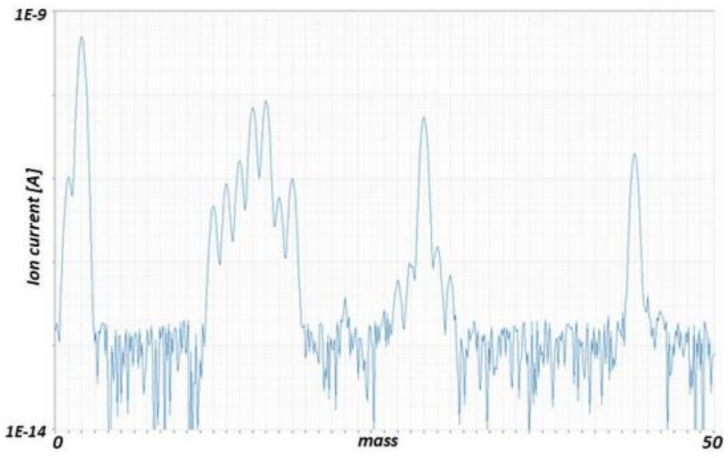
On behalf of the vacuum group

Introduction

- **AI usage in TE-VSC: Limited, with a promising future**
- **Automatic Heating Detection from Vacuum Pressure Readings – Collaboration ABP & VSC**
  - Goal: automatically classify and detect abnormal patterns indicative of temperature increases
  - Paper: Machine learning for beam dynamics studies at the CERN Large Hadron Collider
- **Automated Mass Spectra Analysis - Collaboration with IDAL (University of Valencia)**
  - Goal: Automatic gas identification in mass spectra
- **Generative AI for coding**
  - Tool: Jetbrains AI assistant (ChatGPT-based, paid license)
  - Applications: coaching, generation of code snippets, reviews, automated documentation
- **AI tools are promising for diverse Vacuum related problems including:**
  - Advanced fault detection / diagnostics
  - Data Analysis
  - Predictive maintenance
  - Process simulation



Indication of temperature increase from data correlation



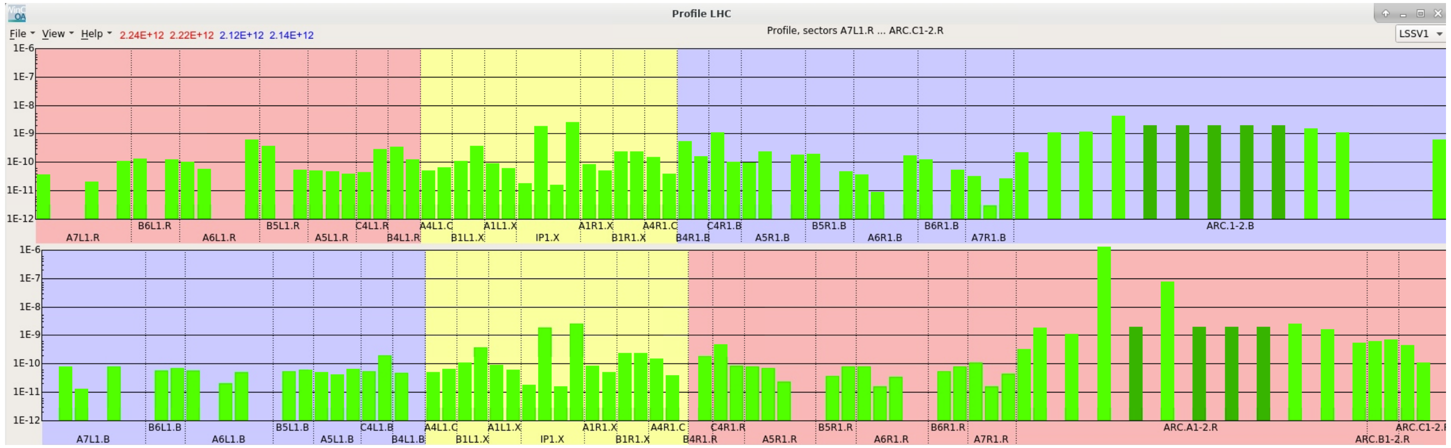
Mass spectrum

Some words of caution

- **No feasibility studies have yet been performed on the cases we’re about to show**
- **Watch out for the golden hammer: over-reliance on a shiny new tool**
- **ML algorithms likely outperform traditional algorithms when:**
  - Handling Large and Complex Data Sets
  - Discovering Non-Linear and Hidden Relationships
  - High Dimensionality (input parameters) problems
  - Continuous Learning and Adaptation
  - Natural Language Processing and Understanding
  - Image and Pattern Recognition

Pressure Anomaly Detection

- **Currently, vacuum levels are compared to fixed pressure baselines (piquet daily monitoring + threshold alarms). If the pressure is below a predefined threshold, it is considered good enough. This is a rather unrefined method:**
  - **Lack of historical context:** Doesn't take into consideration past pressure values, which might have indicated conditions were degrading
  - **Beam characteristics ignored:** Doesn't consider presence or absence of beam and its characteristics which influence pressure
  - **Non-linear pressure dynamics:** multiple phenomena contribute to the pressure and define what's normal or not
  - **Inability to detect early signs of failure:** subtle deviations from normal pressure patterns can go undetected



Pressure Anomaly Detection

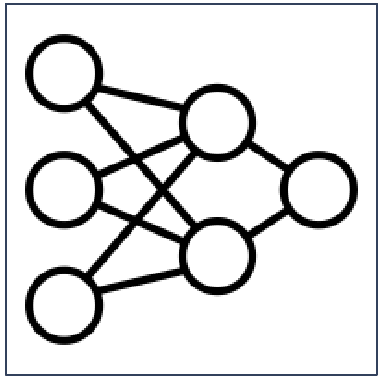
- **A vacuum classifier to provide a deeper insight into pressure evolution in each sector**
- **Causes of anomalies could be identified in a similar way**

Input Parameters

Current Pressure
Reference Pressure
Pressure X mins moving AVG
Pressure Max Value last X mins
Beam Current
Number of bunches
Temperature

→

Anomaly Detection



→

Classification

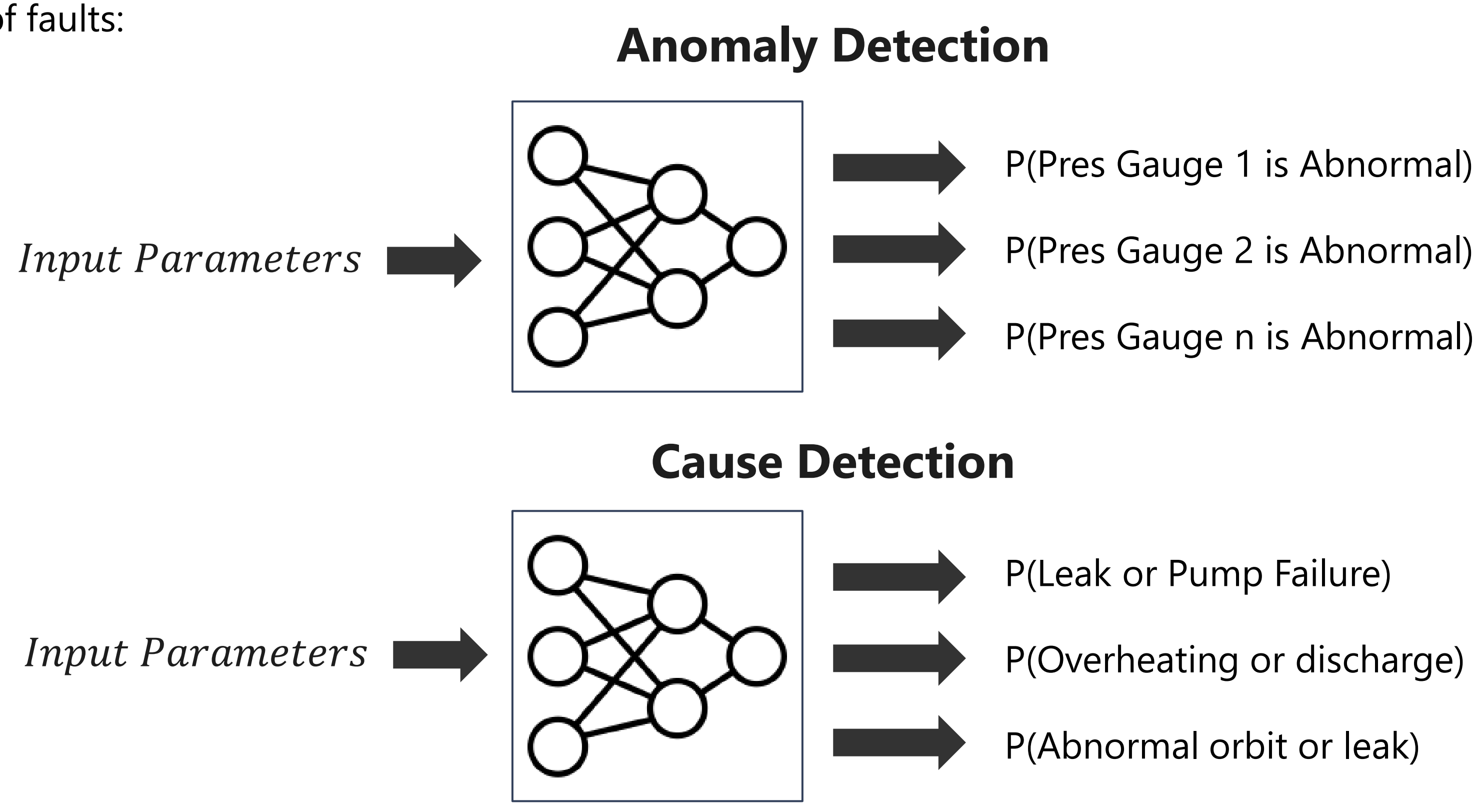
P(Pressure Improving)

P(Pressure Stable)

P(Pressure Degrading)

SuperKEKB Pressure Anomaly detection

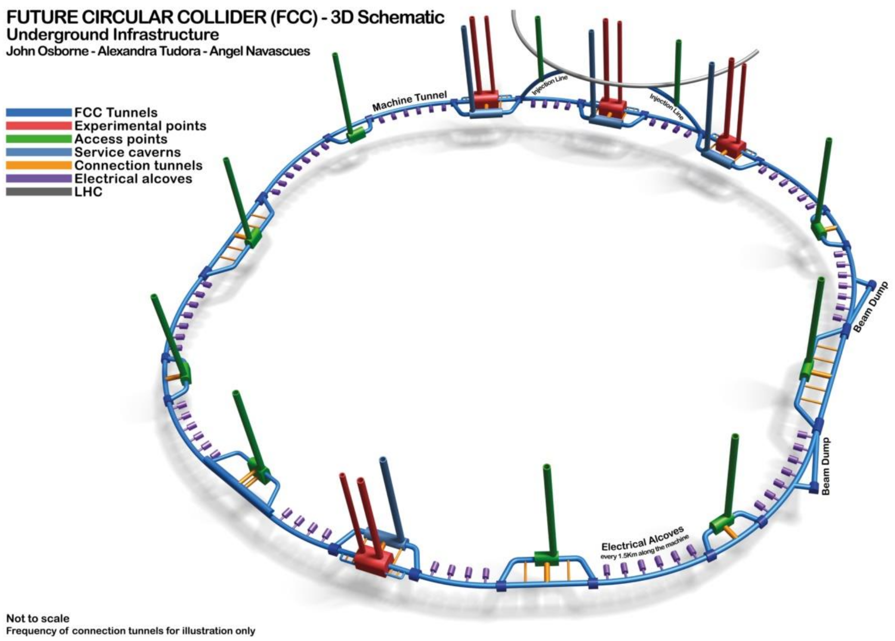
- **Objective:** Develop a pressure anomaly detection system using ML for the SuperKEKB accelerator's vacuum system
- **Fault detection:** detects faults in pressure measured by ~600 gauges in real-time, ensuring early intervention
- **Methodology:** 2 Feed-forward neural networks (FNN) trained on actual abnormal behaviors
- **Fault classification:** can detect the following types of faults:
  - Leak or pump failure
  - Overheating or discharge
  - Pumping down or leak
  - Abnormal orbit or leak
  - Pressure burst or leak



Suetsugu, Y. (2024). Machine-learning-based pressure-anomaly detection system for SuperKEKB accelerator. *Phys. Rev. Accel. Beams, 27*, 063201. DOI: 10.1103/PhysRevAccelBeams.27.063201.

The Future of Vacuum System Monitoring

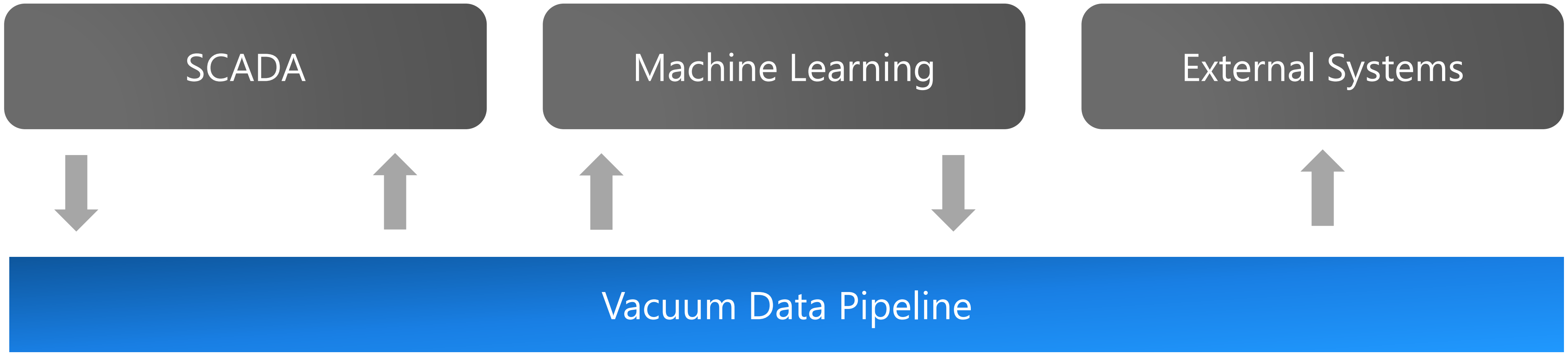
- **Monitoring of large-scale control systems expected to include some forms of machine learning algorithms**
- **Manual monitoring of a vacuum system the size of FCC is impractical**
  - Effective Monitoring of thousands of gauges requires automated tools
- **Early fault detection is critical**
  - Scaling up the number of devices may bring availability issues
  - Identification of issues before they escalate maximizes intervention opportunities
- **Automation is economically essential**
  - The cost of downtime of the FCC due to vacuum system failures most likely justifies investment in automated diagnostics



- **We should start gaining knowledge on machine learning today if we are to deploy mature tools for the FCC**
  - The LHC could be the test bench of FCC in terms of machine learning

Running ML Algorithms in the Vacuum Control System

- **Data Requirements for AI:** running ML algorithms requires processing of vast amounts of data, either in batch or stream
- **Batch processing:** already possible using NxCALS - ideal for offline data analysis where processing delays are acceptable.
- **Stream Processing:** Enables near real-time data processing. To maintain optimal performance of the vacuum SCADA, implement a data pipeline utilizing stream processing engines like Kafka, which ensure scalability and redundancy.

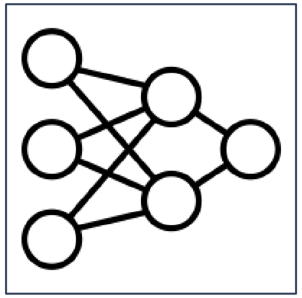


Application Failure Detection

- Vacuum software applications produce over 500k log messages per day
- Important log messages are often hidden within thousands of less relevant entries
- Traditional methods (peak detections, visual search, keywords search are insufficient), especially when multiple events occur simultaneously (eg: failures will be obscured by thousands of high severity messages when PLCs are under maintenance for example)
- What if we could have something looking into every log message and application metric and flag the unordinary events ?

The needle in the haystack

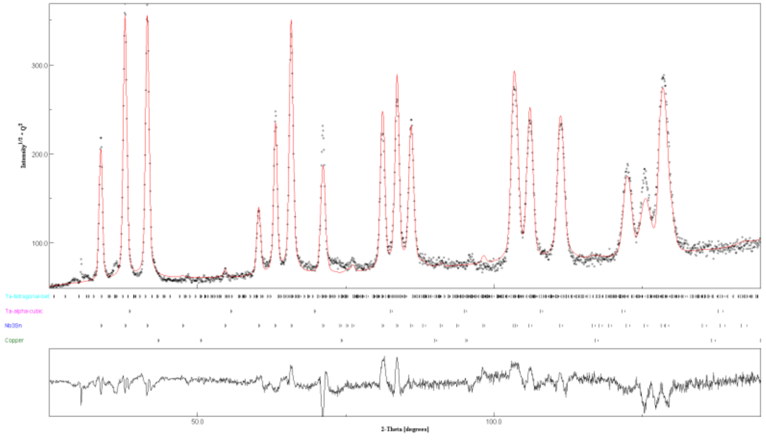
Application Logs
Memory Used
Network latency
Requests per Second



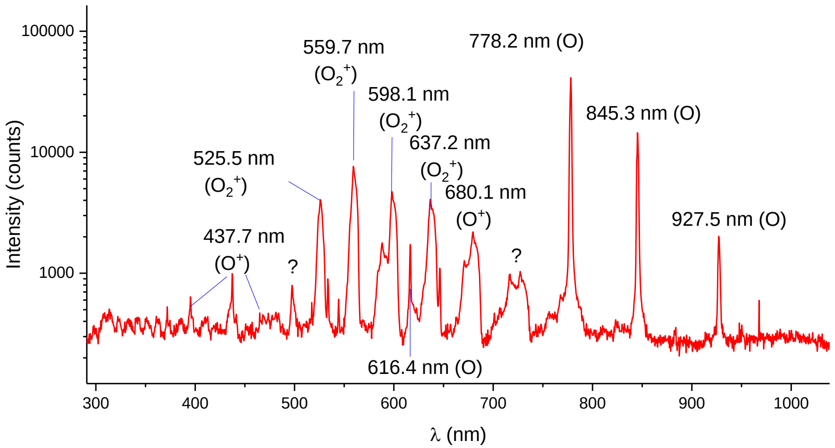
Log Message: SEVERE
“You should really take a look into this one”

Material and Plasma Analysis

- **Purpose:** Analyze materials to identify structure, composition, and reactions
- **Process:** Several techniques involved:
  - **X-ray Diffraction -** crystallographic structure, phase composition, and other structural properties of materials
  - **Optical Emission Spectroscopy** - Optical Emission Spectroscopy (OES) to analyze the light emitted by different species in the plasma, revealing the ongoing chemical processes
  - **Others -** (X-Ray fluorescence, I-V Measurements)
- **Challenge:** manual analysis techniques are quite time consuming
- **ML Objective:** Use machine learning to automate and accelerate the analysis of diffraction patterns and spectra

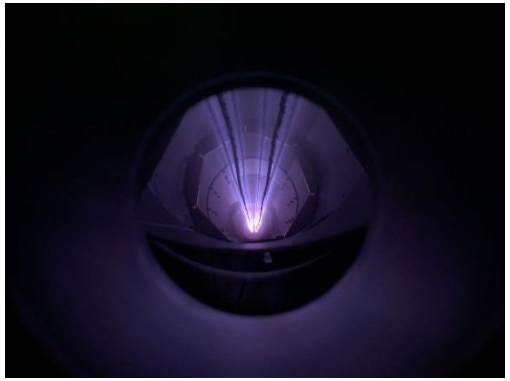


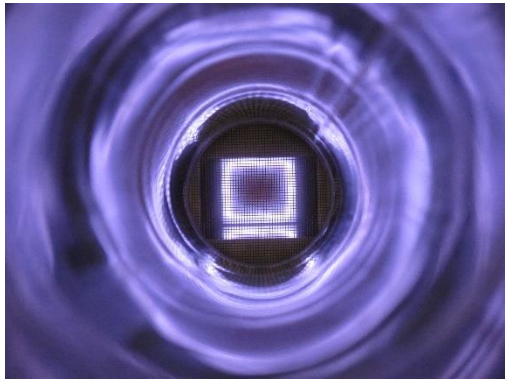
X-Ray Diffraction Pattern



Optical emission spectrum

Plasma Control for Coatings

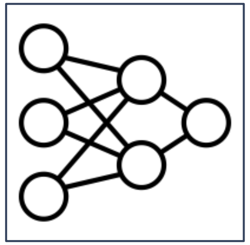




- **Purpose:** Plasmas are used to coat surfaces for various applications (beam pipes, RF cavities, etc)
- **Process:** Coatings are performed by Plasma Control – using High-Power Impulse Magnetron Sputtering (HIPIMS) or Direct Current Magnetron Sputtering (DCMS)
- **ML Objective:** Identify the plasma settings (voltage current frequency, duty cycle, power, pressure) that will optimize one or several intended characteristics of the coating (surface resistance, field dependency, maximum achievable field, external field sensitivity, thermal conductivity, impedance, and roughness.)

Input Parameters

- **Thin Film Characteristics**
- **Optical Emission Spectroscopy**
- **X-Ray Diffraction**
- **X-Ray Fluorescence**
- **Current Voltage Measurements**



Output

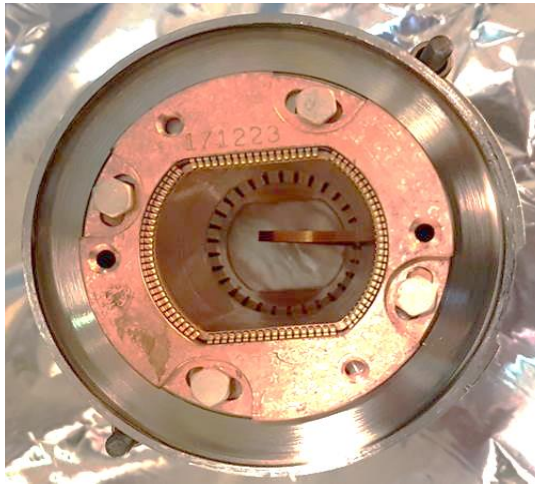
[voltage current frequency, duty cycle, power, pressure, etc..]

Optimized for 1 or more thin film targets (impedance, maximum field, etc)

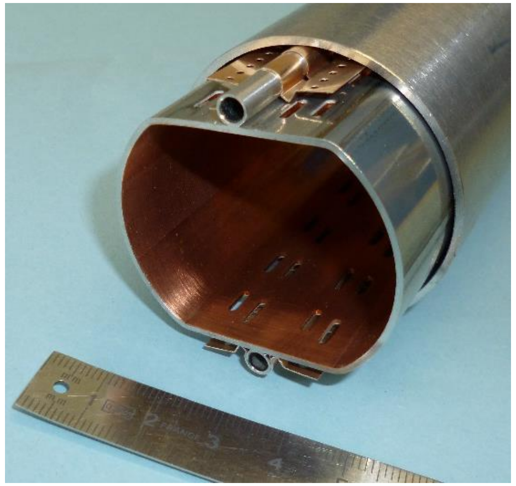
Automated Beam Screen Inspection

- **Purpose:** Inspection system for the LHC beam screen
- **Process:** An Inspection system with cameras will travel an entire ARC (2.8 Kms) to identify blockages in the beam screen
- **Objective for ML**: Automated recognition of beam pipe blockages
- **Challenge**: limited size of the device; Inspection system has to be fully autonomous; computation to be performed inside of the device

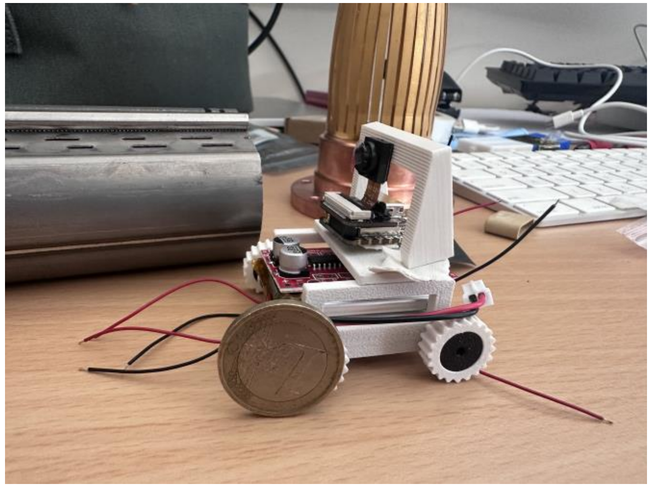
Bent rf finger



LHC Beam Pipe



Prototype Inspection System



Confusion matrix (validation set)

	BACKGROUND	OBSTACLE	PIM	BEAM SCREEN
BACKGROUND	99.8%	0.2%	0%	0%
OBSTACLE	26.3%	73.7%	0%	0%
PIM	0%	0%	100%	0%
BEAM SCREEN	0%	0%	0%	100%
F1 SCORE	1.00	0.65	1.00	1.00

Final Remarks

- **Machine Learning of Stand-alone applications (independent on vacuum controls data)**
  - Can start right away provided resources are available as there are no pre-requisites
  - Examples: X-Ray Diffraction Analysis, Thin Films Coating Optimizations, Beam screen defect analysis
- **Machine Learning on vacuum control system data**
  - **Batch processing:** can already be performed with limited data (eg: pressures / ion pump currents)
  - **Stream processing:** Building a vacuum data pipeline will allow for near real-time processing
    - Ballpark estimate: 1 to 2 FTE Years
- **Current staffing levels in software team are limited**
  - The team is currently focused on maintenance and hardware-related developments necessary for HI-LUMI -> no capacity for R&D
  - Limited in-house knowledge and experience in machine learning
- **Strategic focus on machine learning**
  - To realize the potential of machine learning, a gradual increase in the allocation of time and resources to ML initiatives will be essential
  - Creating strategic collaborations with expert institutes or other groups at CERN is an attractive strategy

Thank you !