Overview of Data Science Dept. Research at Jefferson Lab

Dr. Malachi Schram

On behalf of the Data Science Department

Data Science at Jefferson Lab

Mission:

- Provide world-class data science solutions to advance research in nuclear physics by working with the subject matter experts at Jefferson Lab, partnering Universities and Labs, and the Department of Energy.
- Provide world-class data science solutions to scientific applications relevant to the regional scientific community

Vision:

- Expand the capability and capacity of data science at JLab
- Create a collaborative data science research hub to:
  1. Provide world-class solutions to scientific challenges
  2. Provide real-time optimization solutions for complex system
  3. Champion education and research opportunities with regional Universities and industry
  4. Reduce the carbon footprint by optimizing the data science workflow and algorithms

FY22 LAB S&T Agenda, Milestones – Data Science & AI/ML

Overall Goals:

- Establish JLab as the data science hub for nuclear physics and regional scientific applications
- Develop a core capability targeting key elements of the Basic Research Needs (BRNs) defined in the SciML report, etc.
- Provide education and research opportunities to regional universities and industries

Yearly Activities and Milestones:

- Participate in HEP, NP, ASCR, and SciDAC proposal calls (expand as opportunities become available)
- Participate in DOE data science related workshops and BRN reports
- Continuously reevaluate based on State of the Art (SOTA)
- Expand scientific applications and collaborations

FY22:

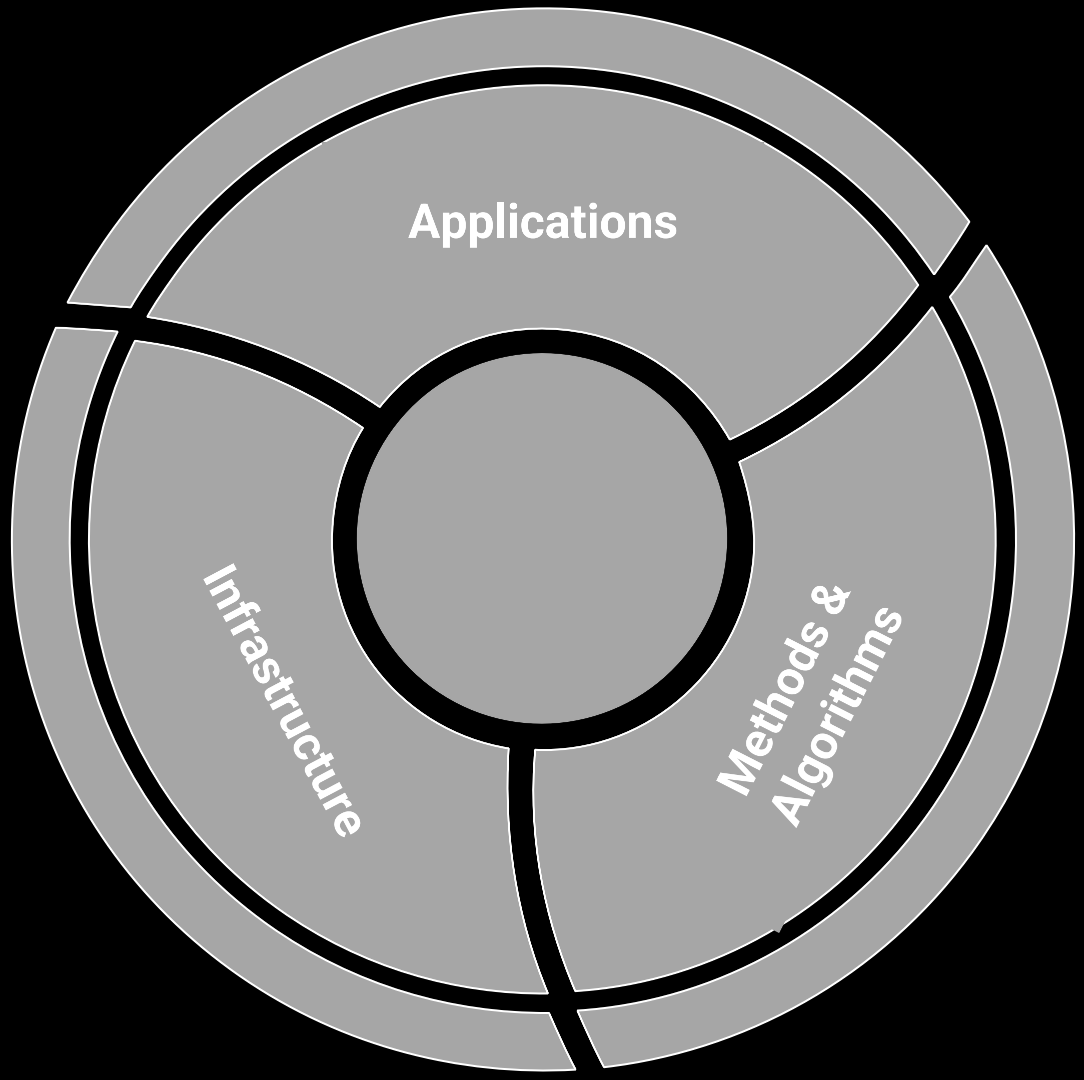
- **Establish a data science priority research needs and data availability for JLab:**
  - **Experimental Halls, Theory, Accelerator, and Facilities**
- Establish collaborations with regional universities and national laboratories
- Develop core capabilities:
  - *Infrastructure*:
    - Evaluate latest SOTA workflow tools for data science
    - Evaluate datasets and model repositories
    - Identify/evaluate needs for digital twin
  - *Methods & Algorithms (addressing the needs of the SciML BRN):*
    - Expand capability in ML-based ***uncertainty quantification*** techniques
    - Develop ***interpretability*** techniques
    - Expand optimization research for ***design and control***
    - Domain-aware ML

Data Science Pillars

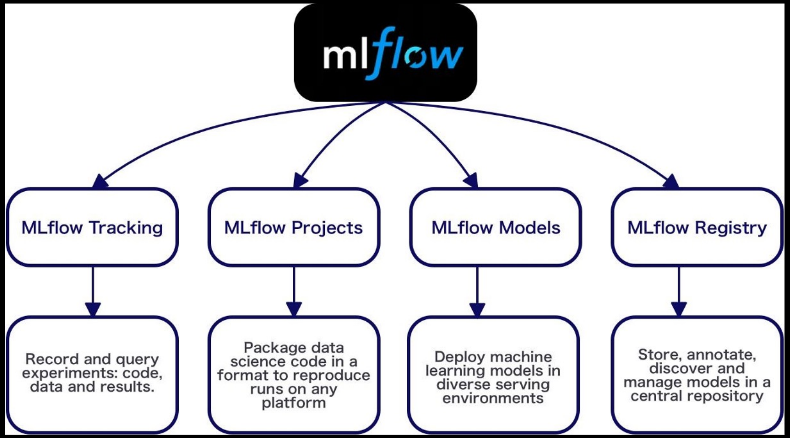
- <u>**Applications**</u>**:**
  - **NP physics**
  - **Advanced computing**
  - **Health & Climate**

- <u>**Focued Methods & Algorithms**</u>**:**
  - **Uncertainty Quantification**
  - **Interpretability and Explainability**
  - **Design & Control**

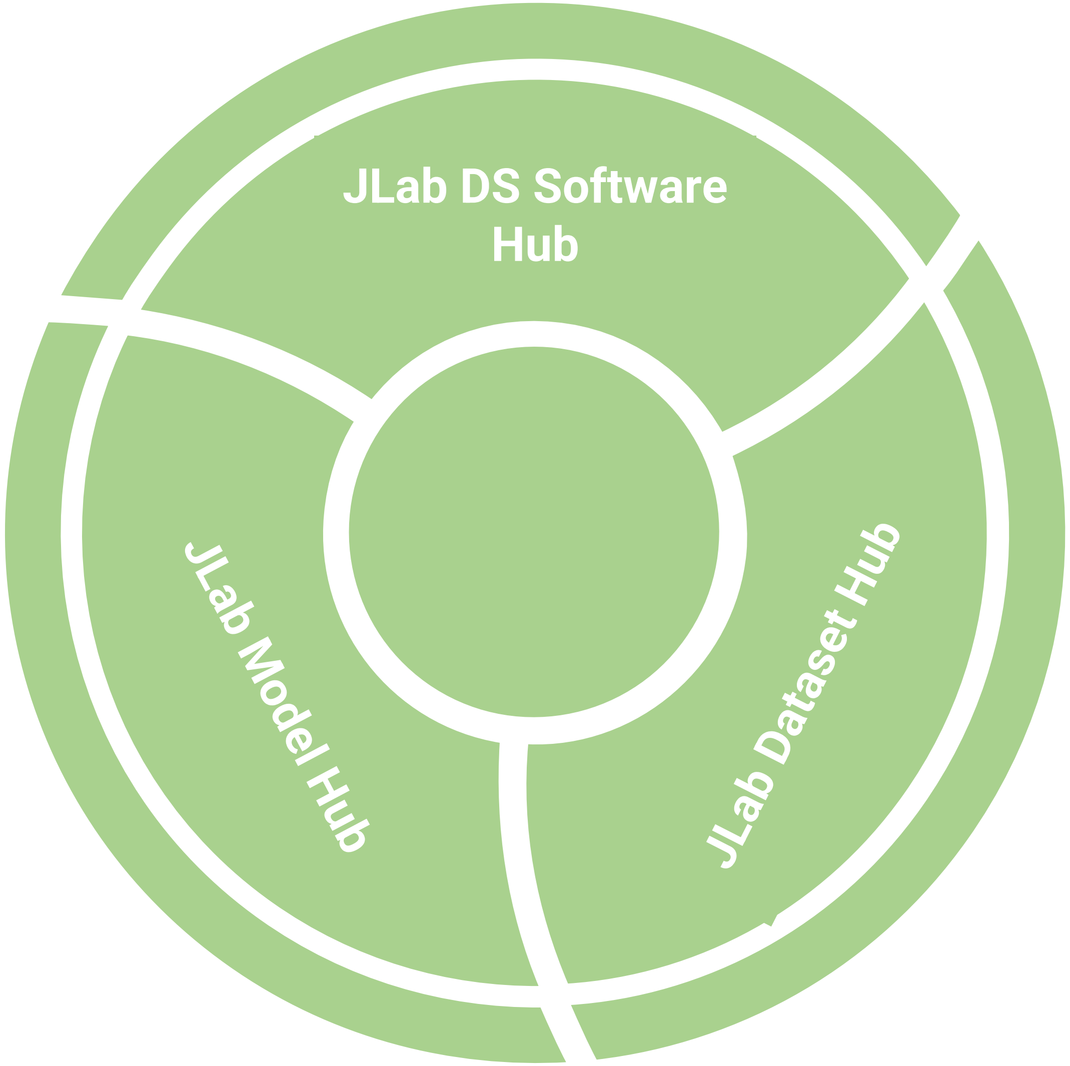
- <u>**Infrastructure:**</u>
  - **JLab ML Hub**
  - **JLab Data Hub**
  - **JLab Data Science software**



Data Science Infrastructure

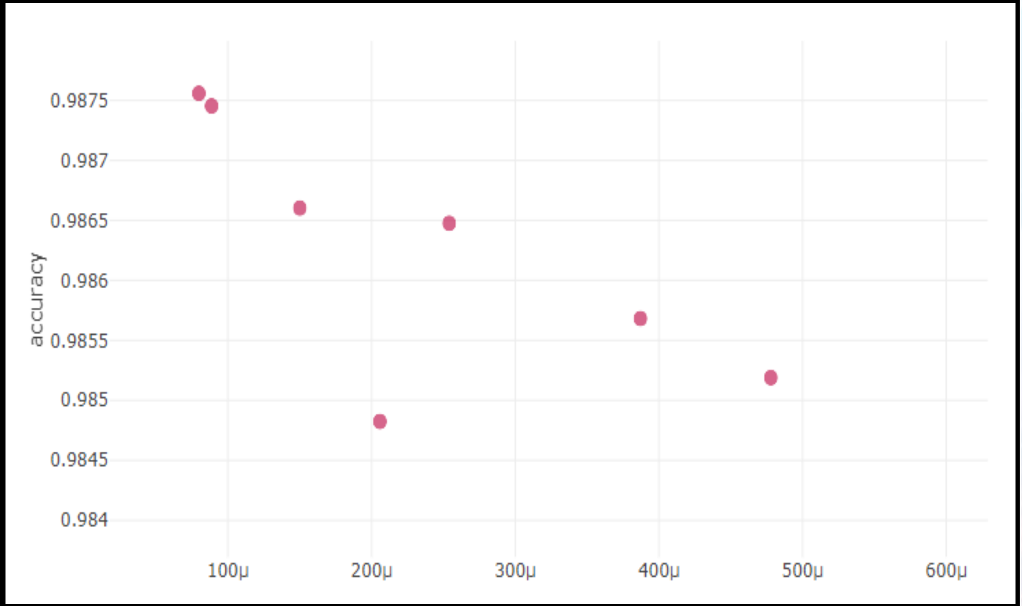


INVENIO
Powering Open Science



JLab ML Hub

- Developing a service for model optimization and curation
- Leveraging and extending MLFLOW
- Support machine learning "**experiments**":
  - Hyperparameter search
  - Model comparison (visualization, etc.)
- Store **Models** for reuse
  - Centralized model registry and repository
  - Versions
  - Searchable by model characteristics and requirements
- Extension captures JLab specific domain metadata and access control policies



Accuracy by learning rate

Data Science Methods & Algorithms

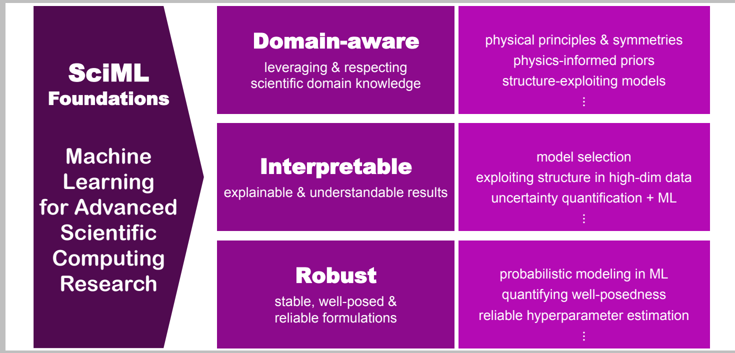


Figure 1: Foundational research themes of SciML must tackle the challenges of creating domain-aware, interpretable, and robust ML formulations, methods, and algorithms.

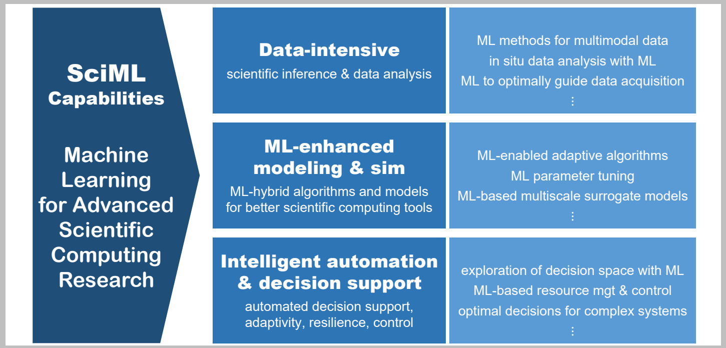
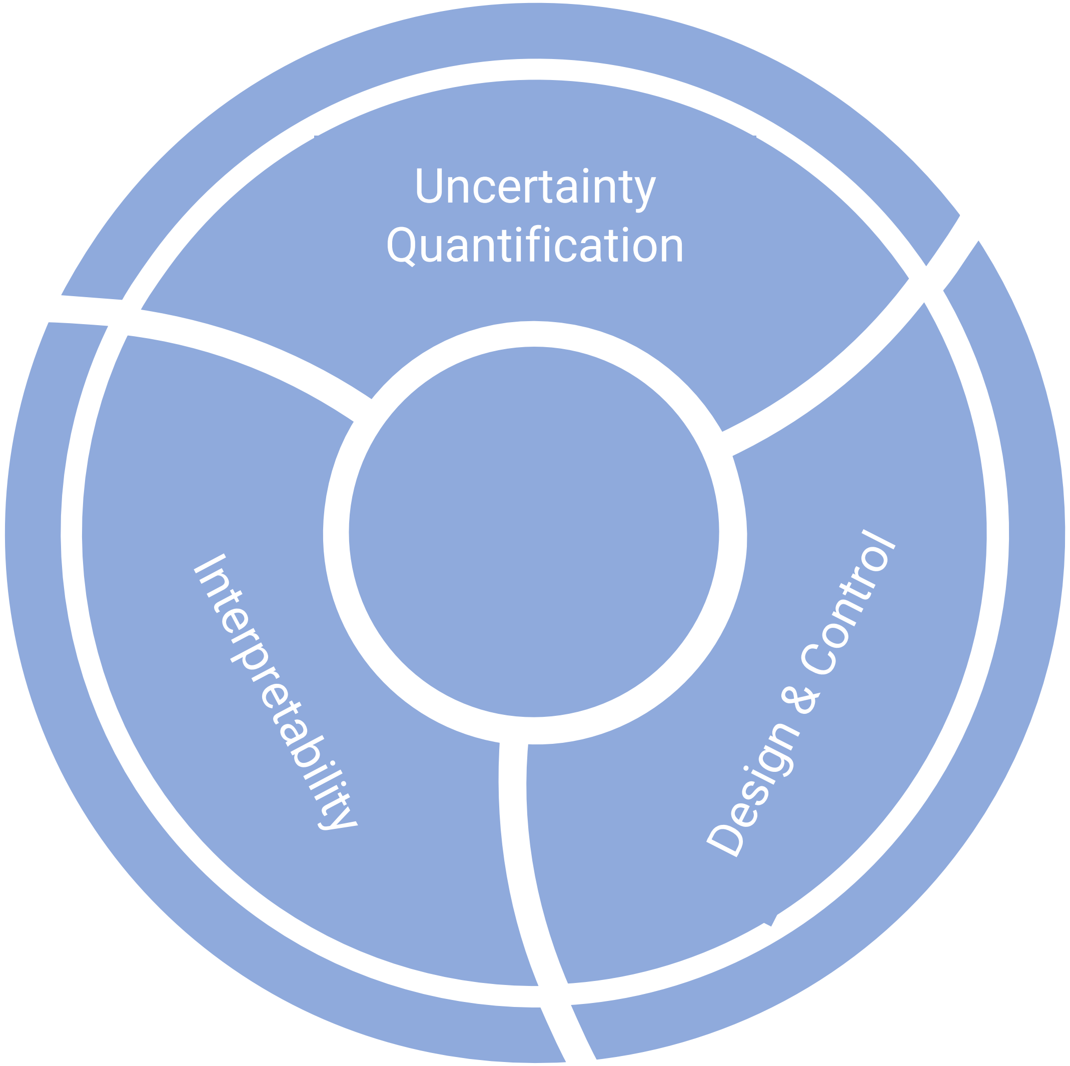
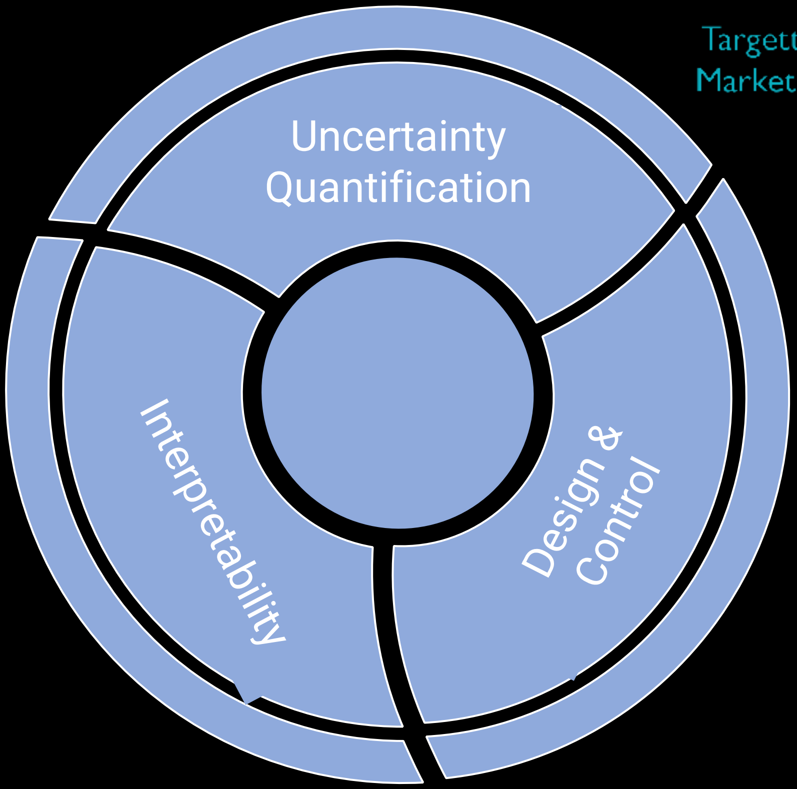


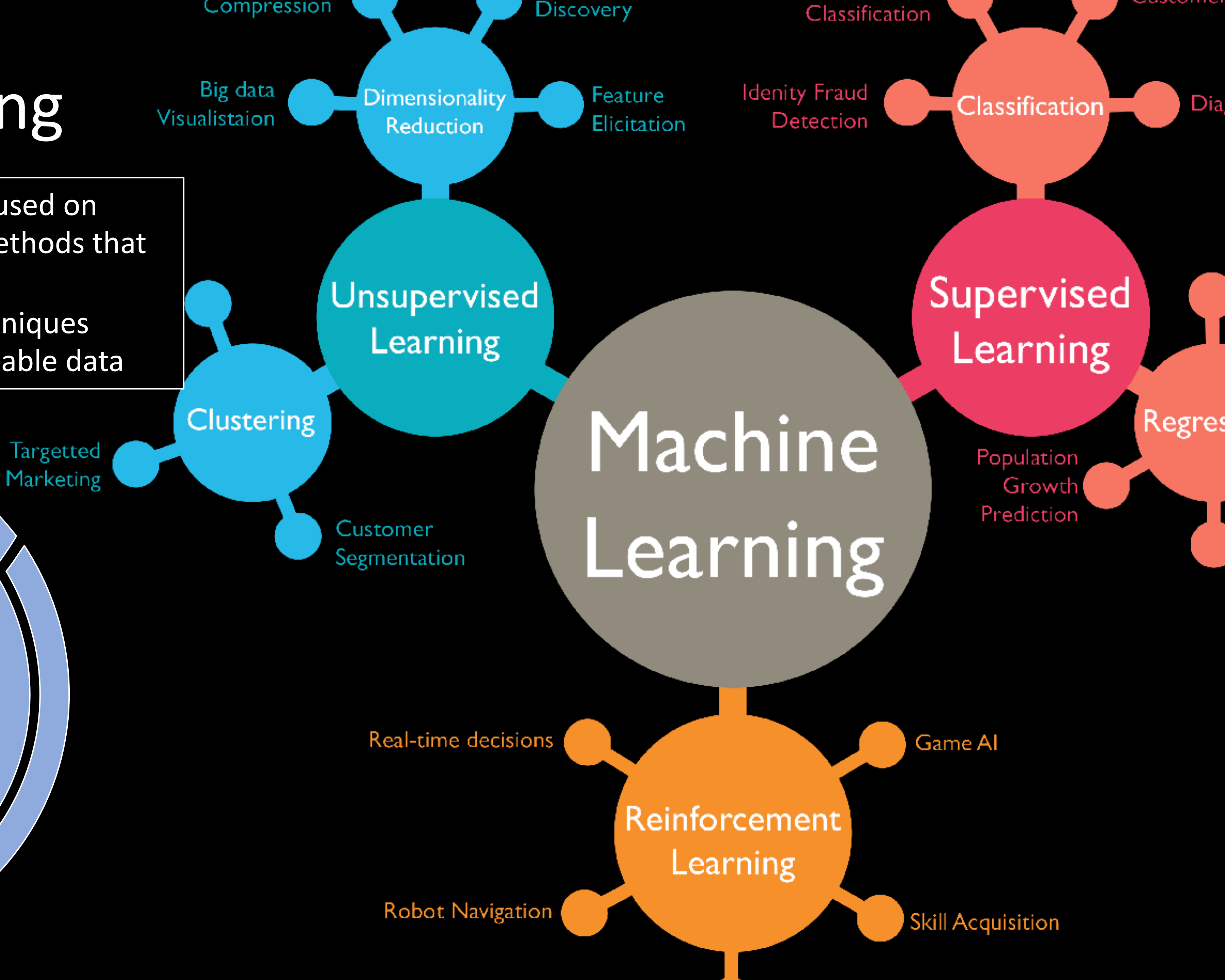
Figure 2: Opportunities for SciML impact arise in scientific inference and data analysis; in ML-enhanced modeling and simulation; in intelligent automation and decision support; and in related applications.



Machine Learning

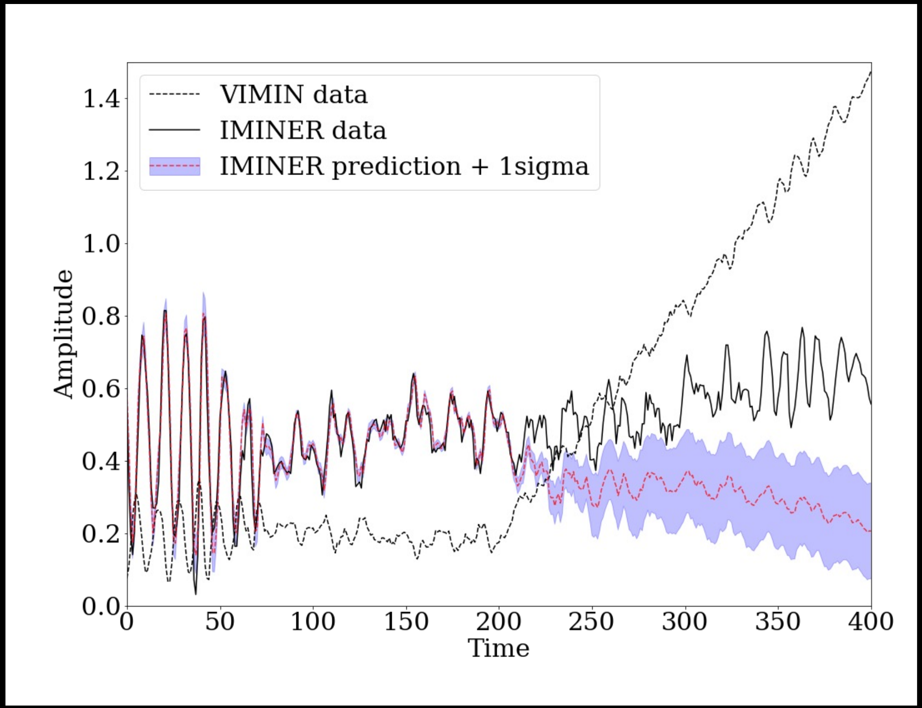
- Machine learning is a field focused on understanding and building methods that 'learn' a set of tasks
- There are a wide range of techniques
- Selection depends on the available data

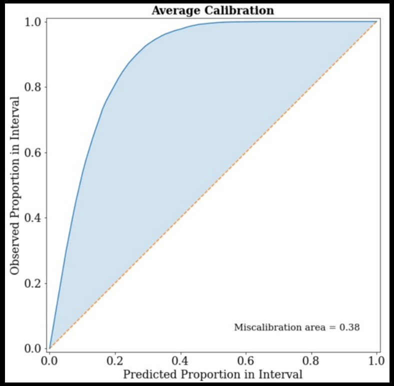


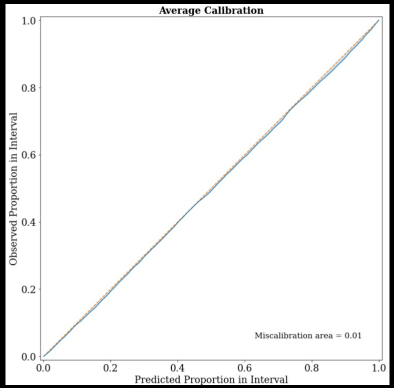


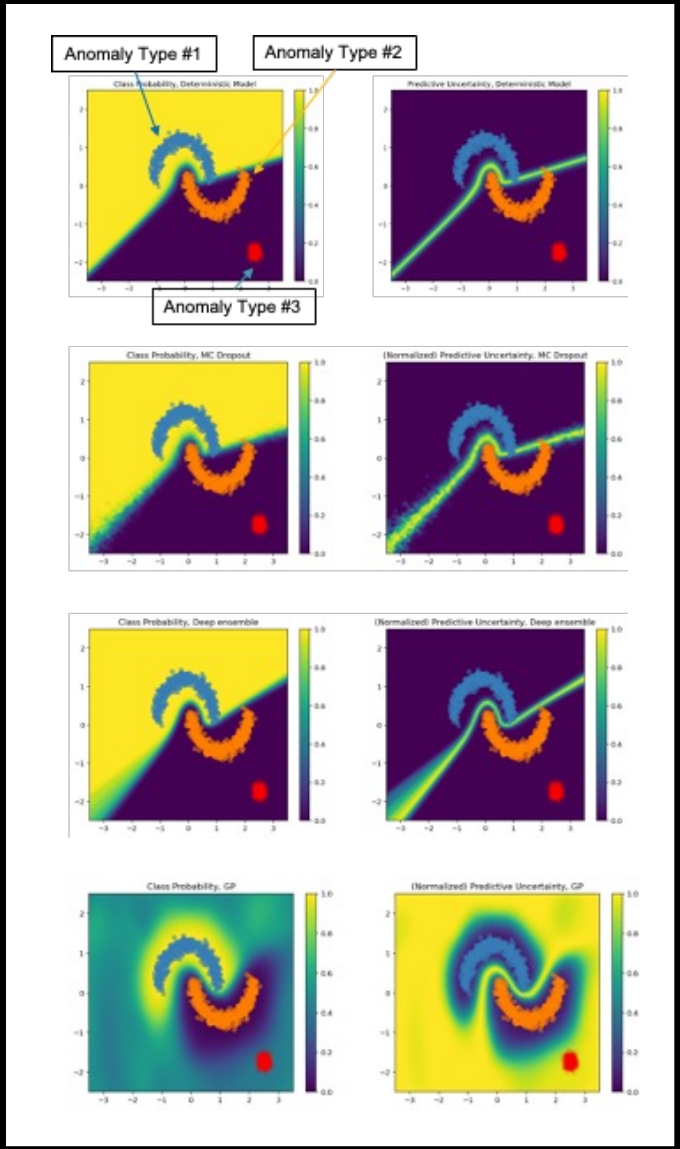
Uncertainty Quantification

- Understanding how to include UQ in deep ML models
- Critical to include out-of-distribution uncertainties
- Incorporate auto-calibration



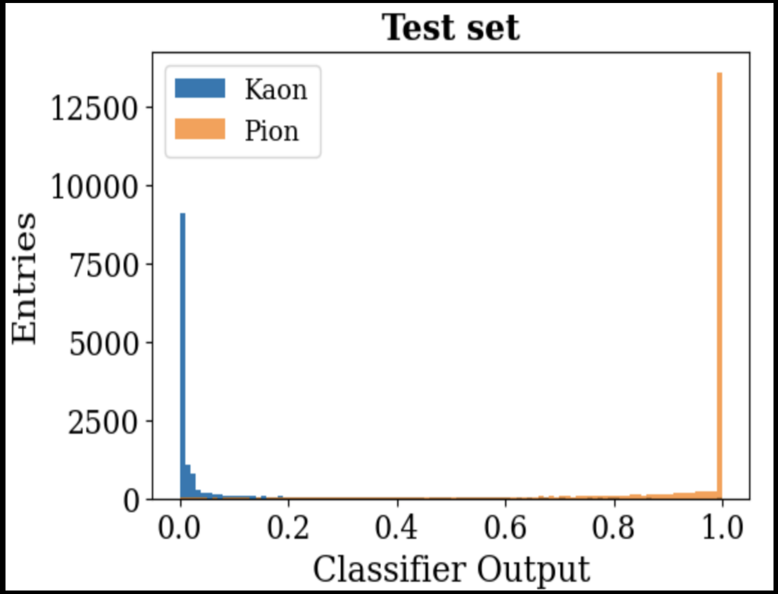


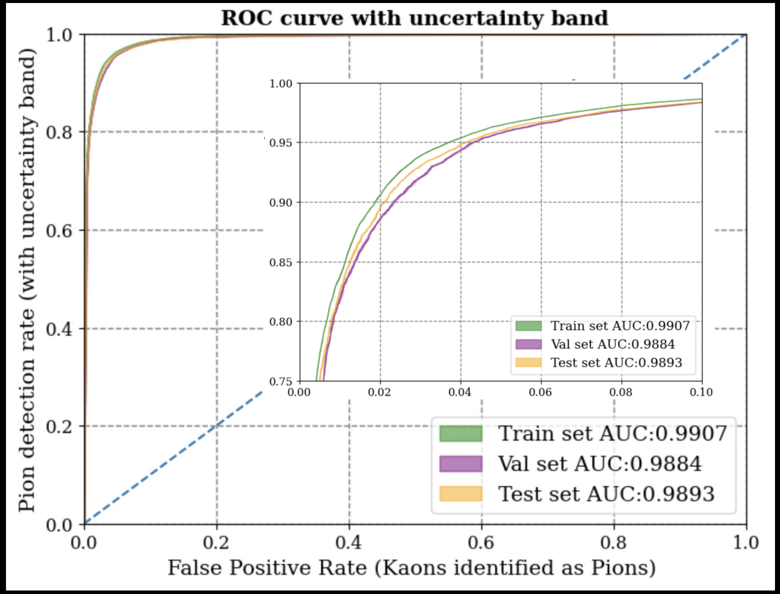


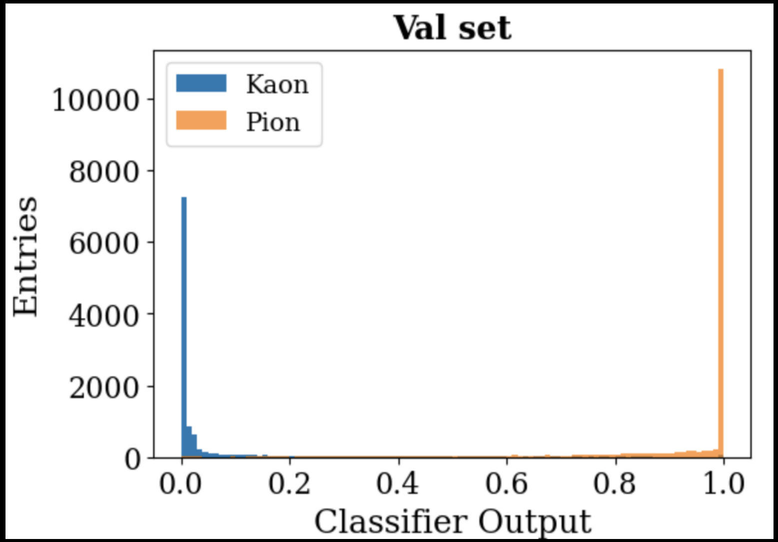


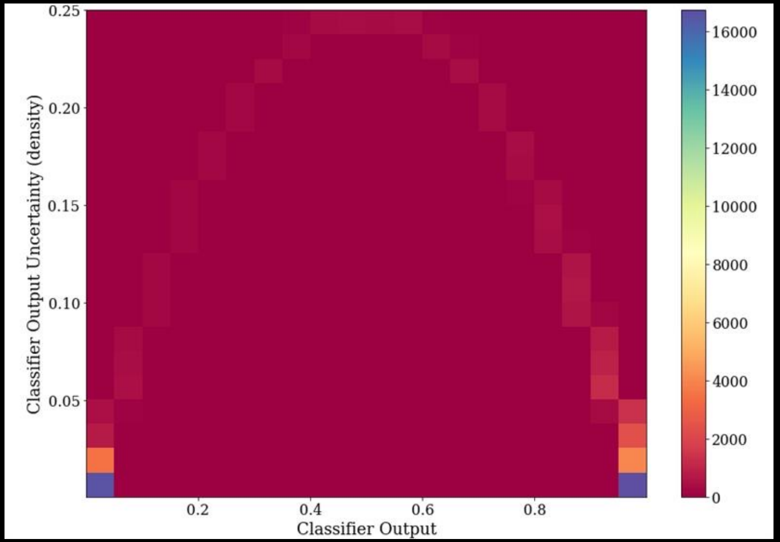
Particle identification for SoLID with uncertainty quantification

- The initial goal was to achieve >95% pion efficiency while keeping false positive below 5%
- Distance aware model provides uncertainty values associated with each output
- We are able to achieve the initial goal on the most difficult kinematic and smallest resolution readout
- With increase in the resolution using other readouts, we believe it is possible to further improve the efficiency
- Performing hyper parameter optimization can potentially improve the accuracy of the model



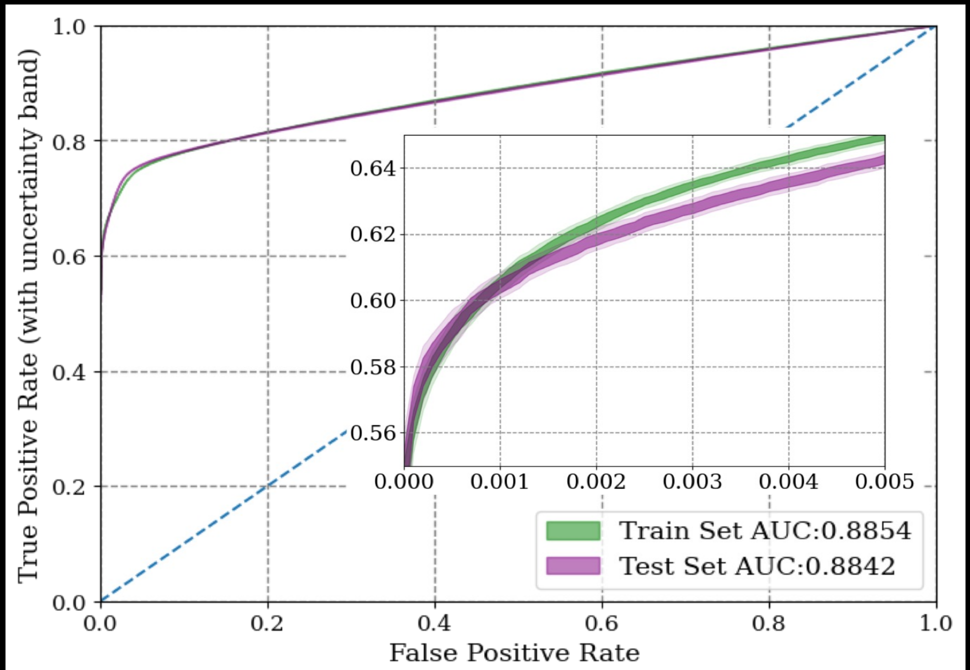


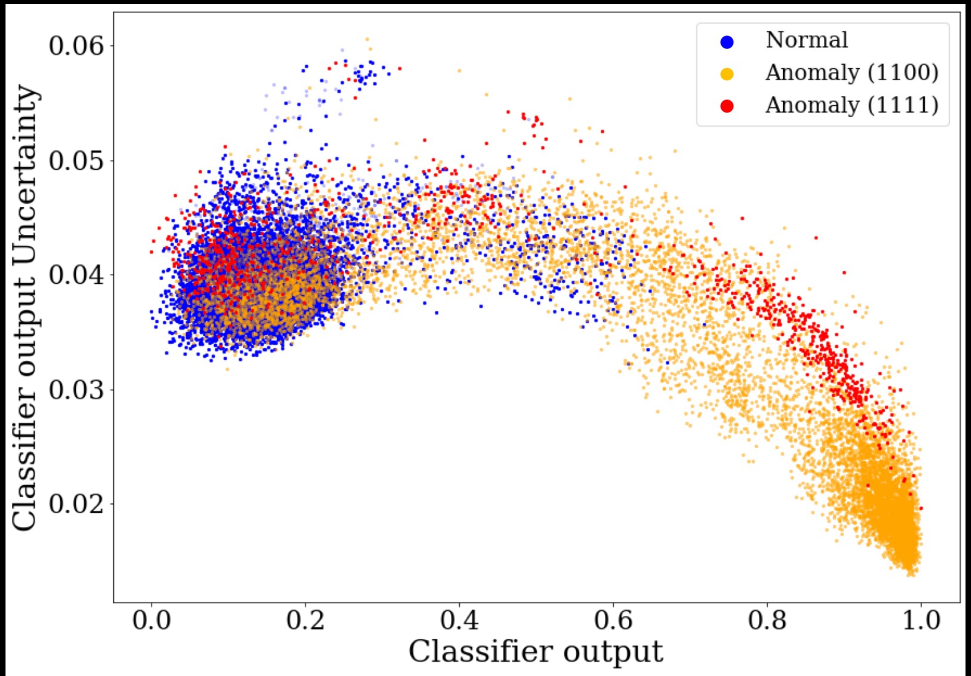




Uncertainty quantification for accelerator anomaly detection at SNS ORNL

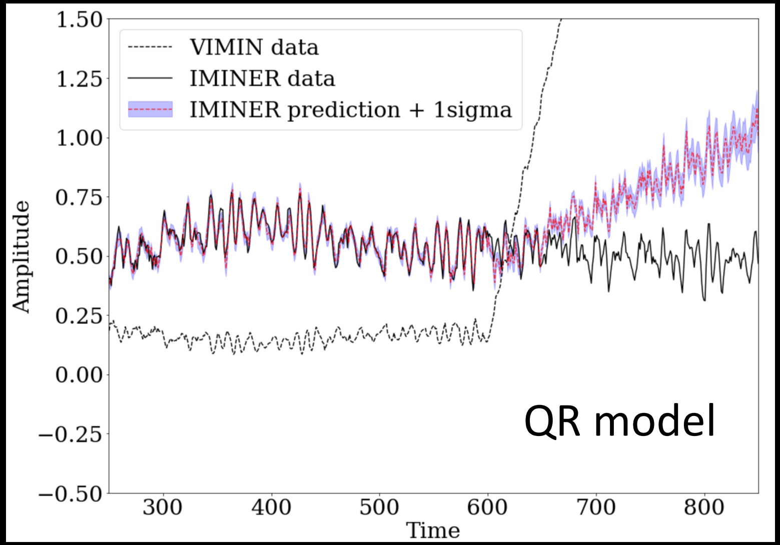
- Results from similarity model showed a ~4x improvement in performance over previously published results
- The ROC curve shows nearly the same level of performance (not optimized)
- We introduced an **out-of-domain anomaly**, labelled 1111 (red), the UQ-based model correctly identified the anomaly and indicated high uncertainty.

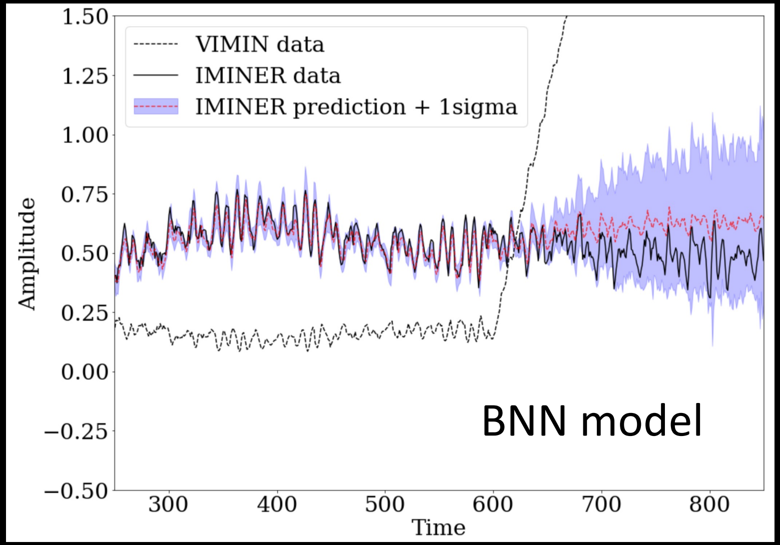


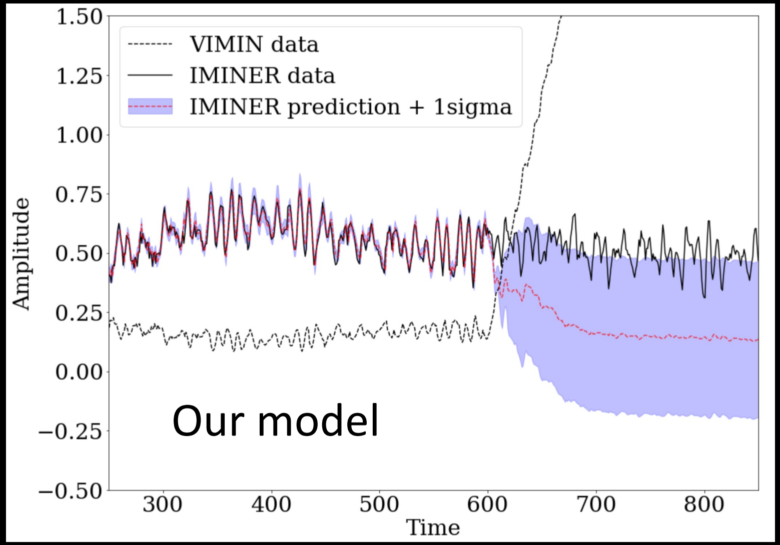


Uncertainty quantification for surrogate models in risk averse control research

- Quantile regression method have great performance in the training distribution and are calibrated by definition, however, the do not perform well for OOD estimation
- BNN models provides does a better job to estimate OOD but require calibration
- GP approximation model provide the best OOD estimation and is calibrated by design

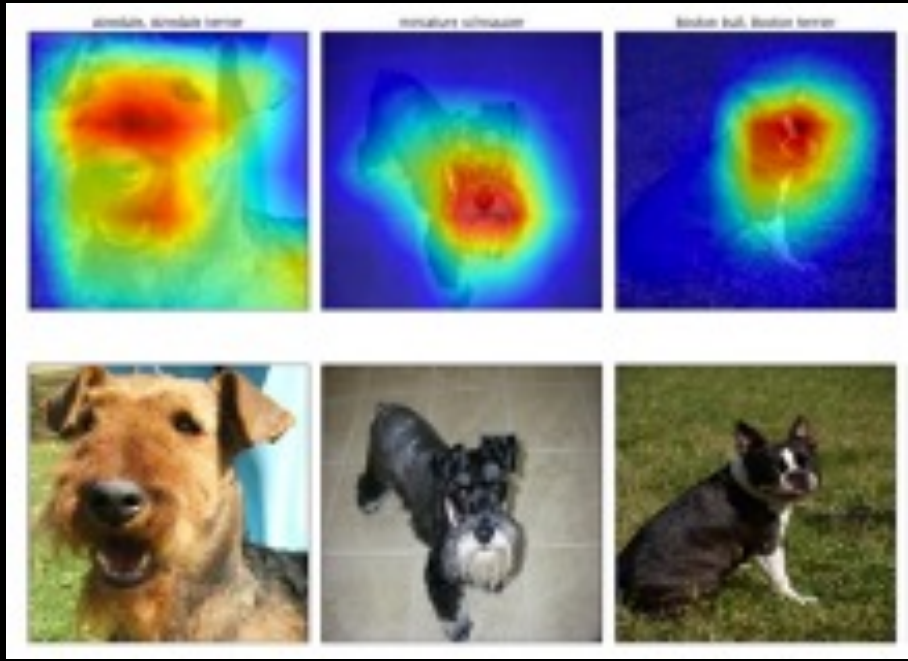




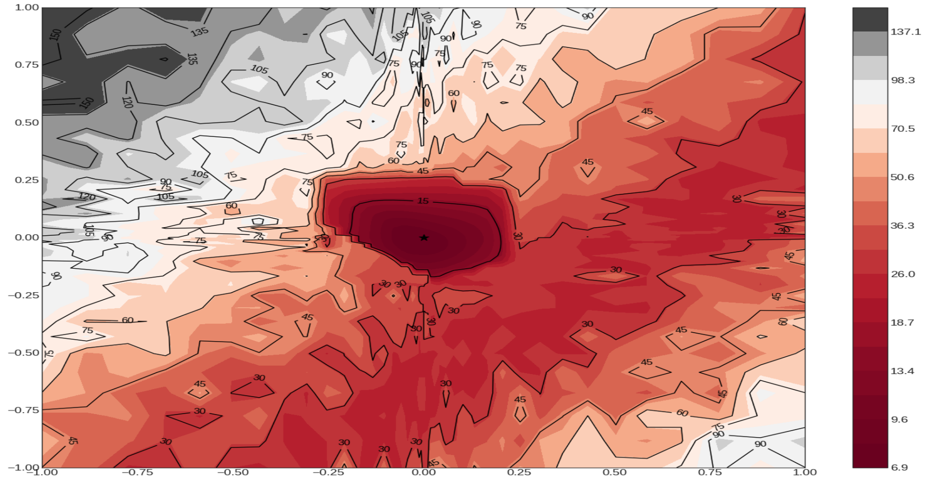


Interpretability, Explainability, and Robustness

- Applying and developing techniques to better understand model predictions and stability
- Gradient activation studies to understand what the model is focusing on
- Loss landscape analysis to better understand the model stability



Loss Landscape for FNAL system dynamic model



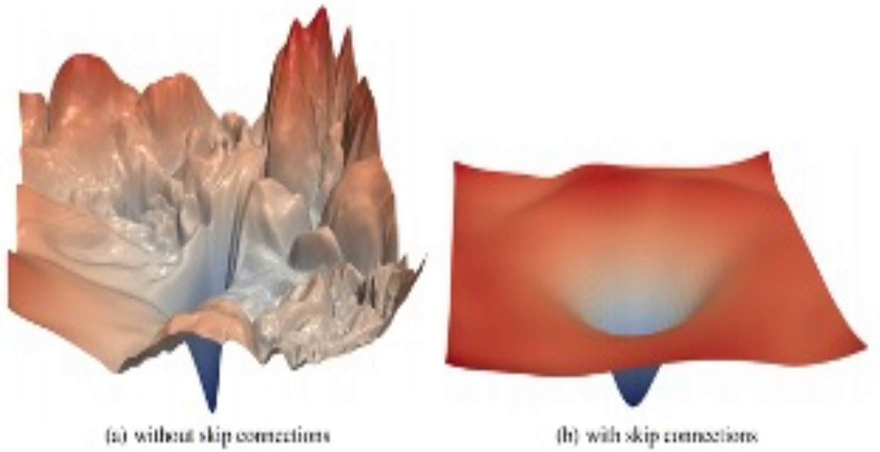
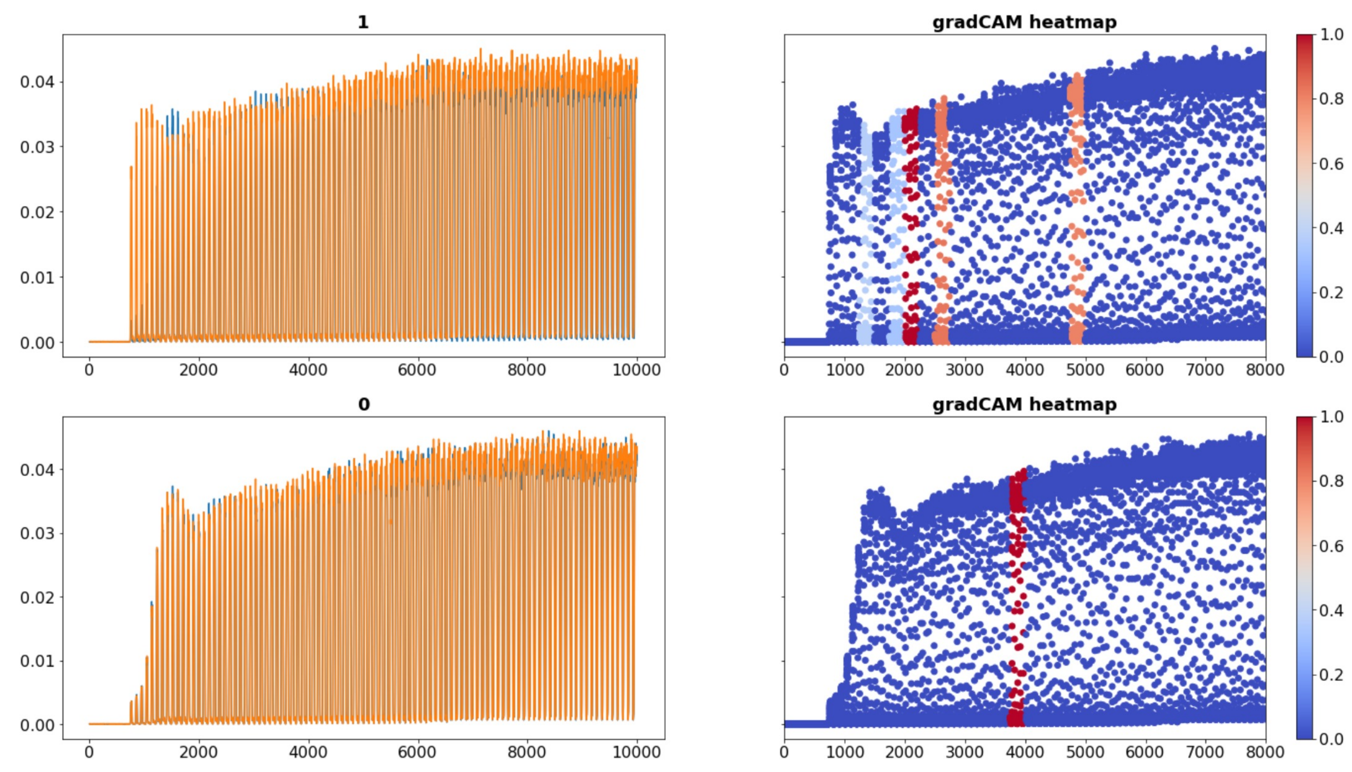
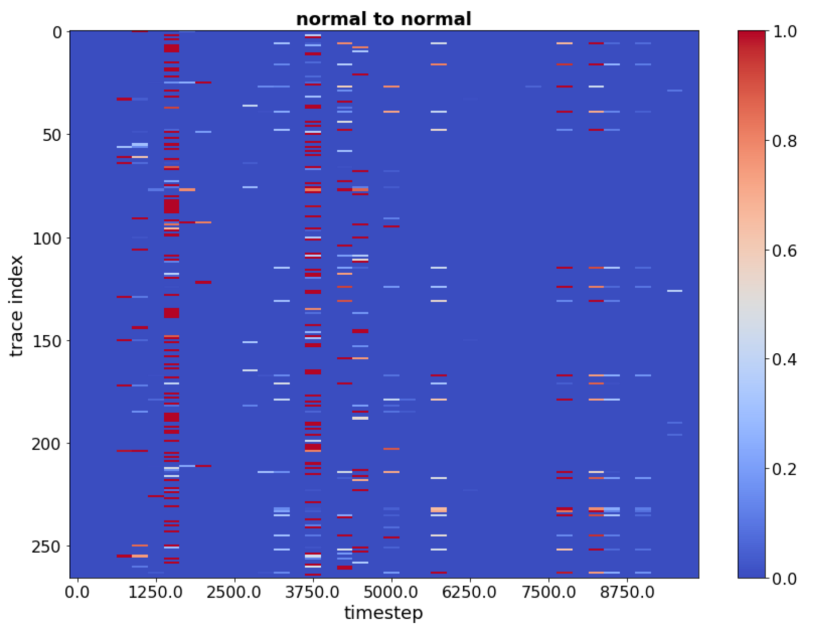


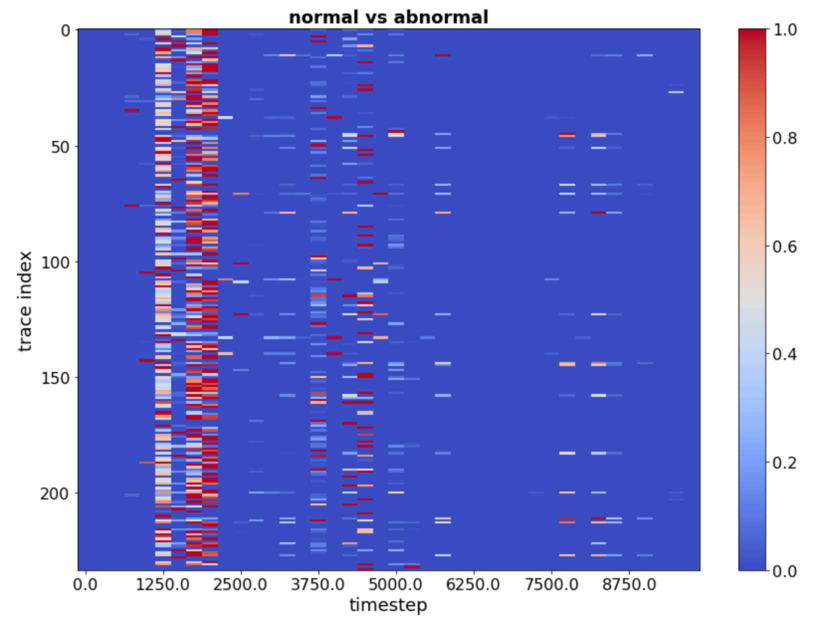
Figure 1: The loss surfaces of ResNet-56 with/without skip connections. The proposed filter normalization scheme is used to enable comparisons of sharpness/flatness between the two figures.

Application at SNS ORNL

- Applied GradCAM analysis on trained ML-based model for errant beam prediction
- Identifies sections of the waveform most relevant for a particular decision from the model
- The results can potentially identified fault types by exploring results

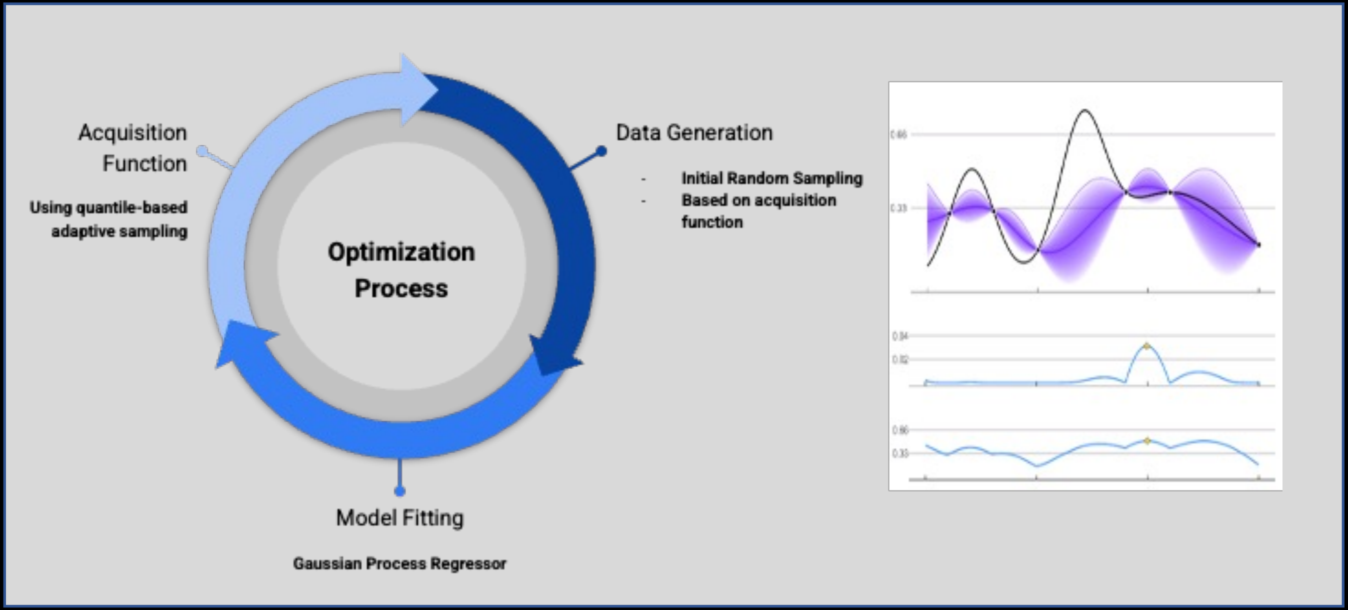


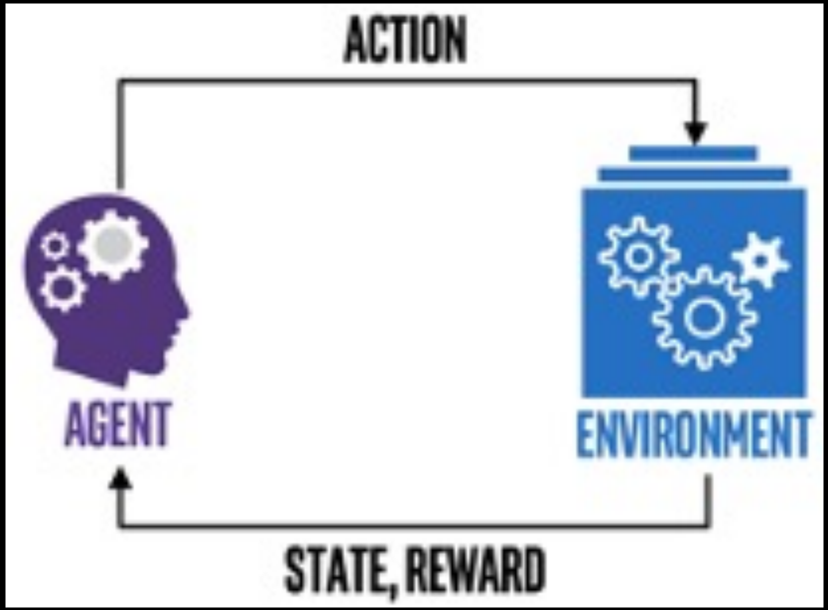


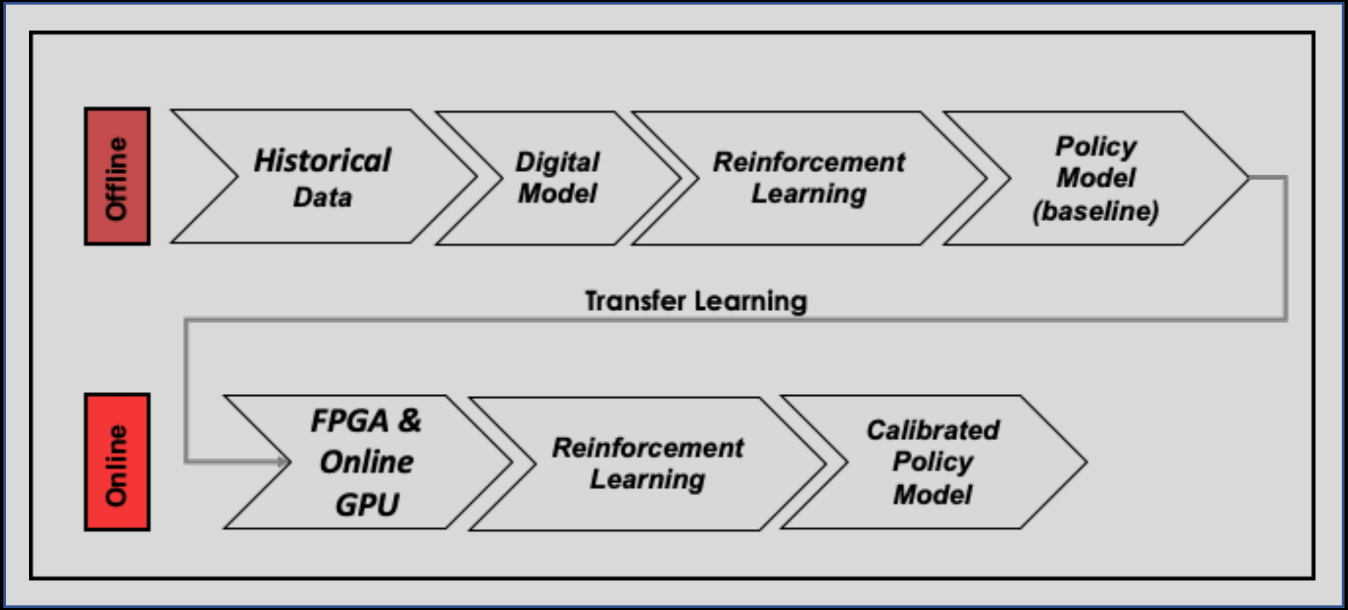


Design & Control

- Advance applications for design & control
- Bayesian Optimization
- Risk Averse and UQ aware Reinforcement Learning







Near real-time control and calibration for the GlueX Central Drift Chamber

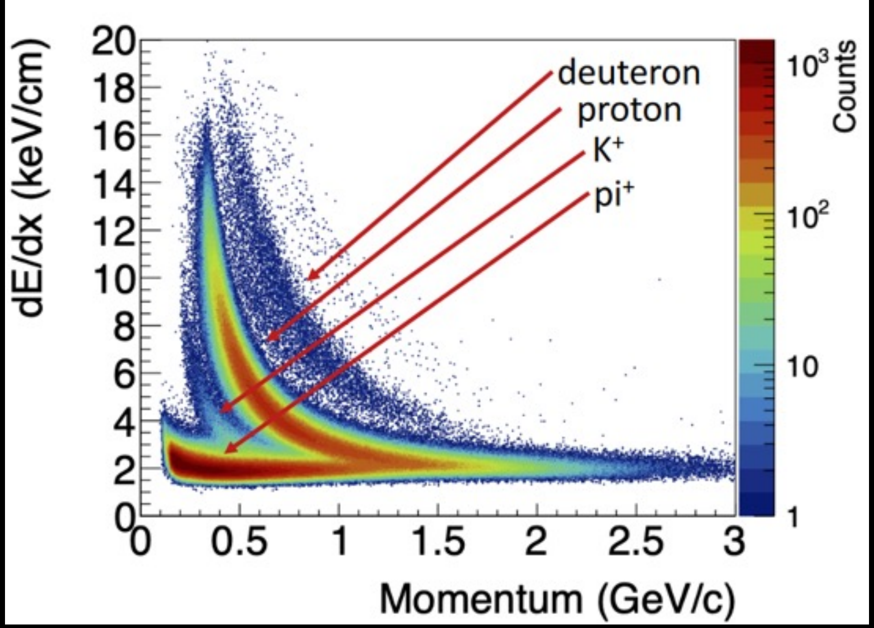
Accelerate the calibration from month(s) to minute(s).

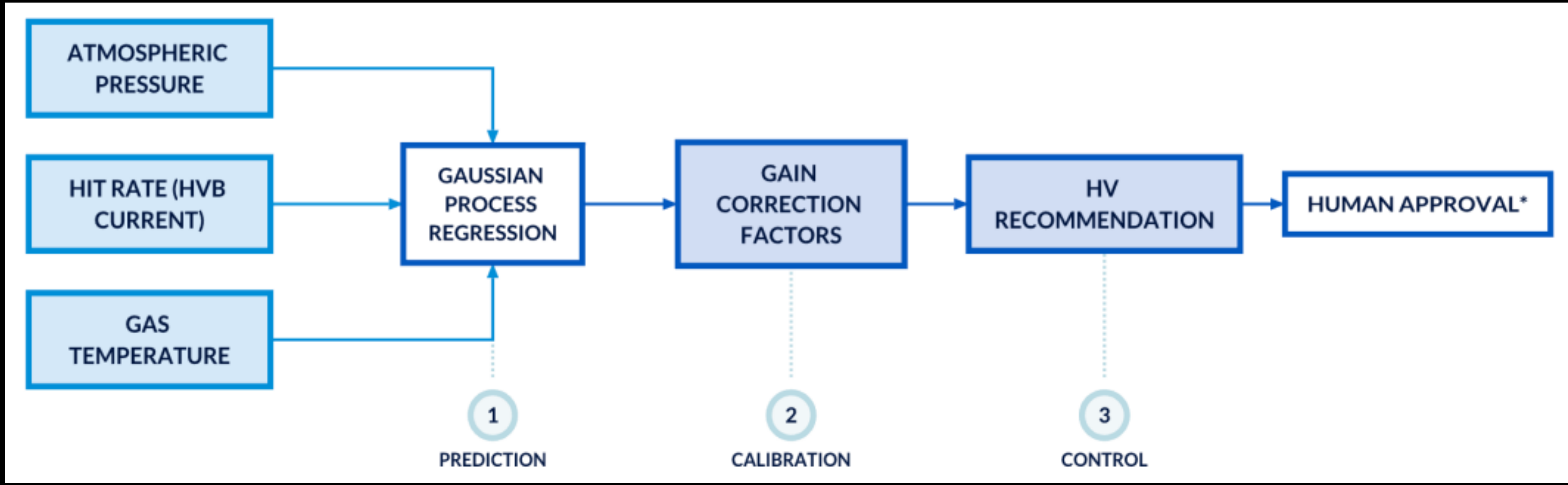
1. Gain Correction Factor: CDC Voltage Gain calibration
2. Time to Distance: track fitting calibration

Calibration is required to provide reliable PID for physics analysis

Considerations:

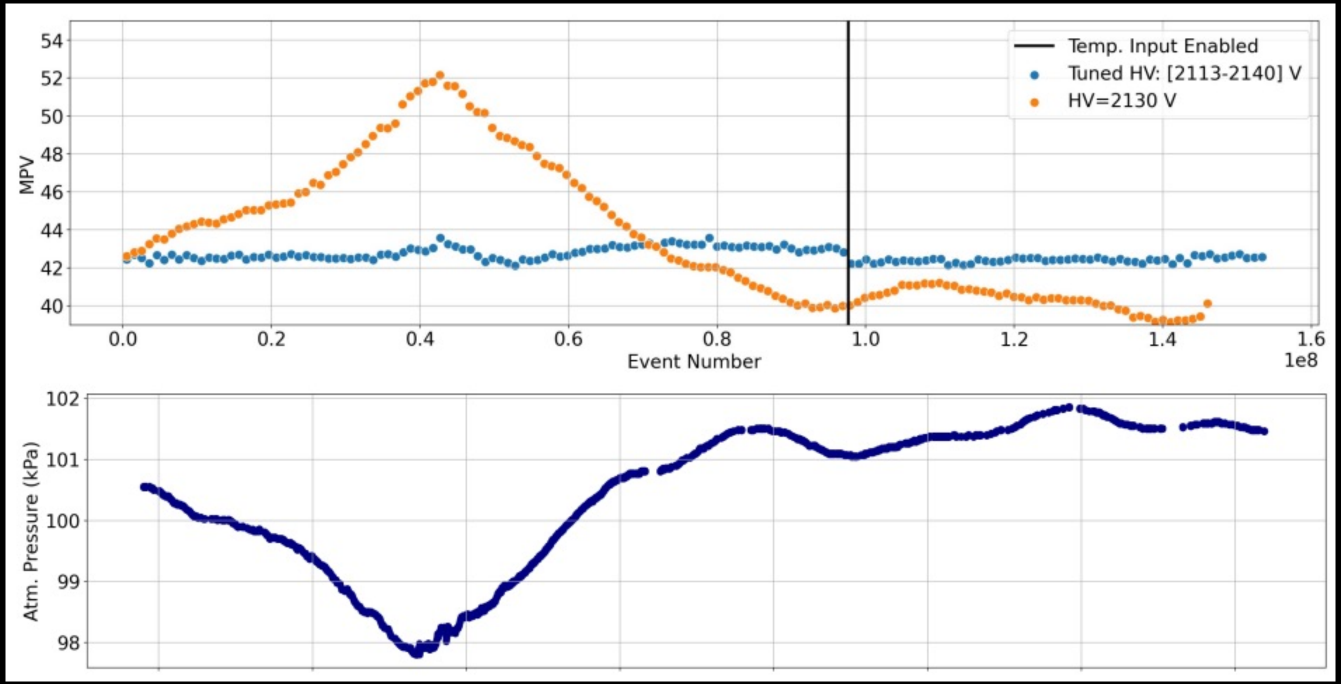
1. External environmental conditions (temperature, pressure)
2. Changing beam conditions (current)

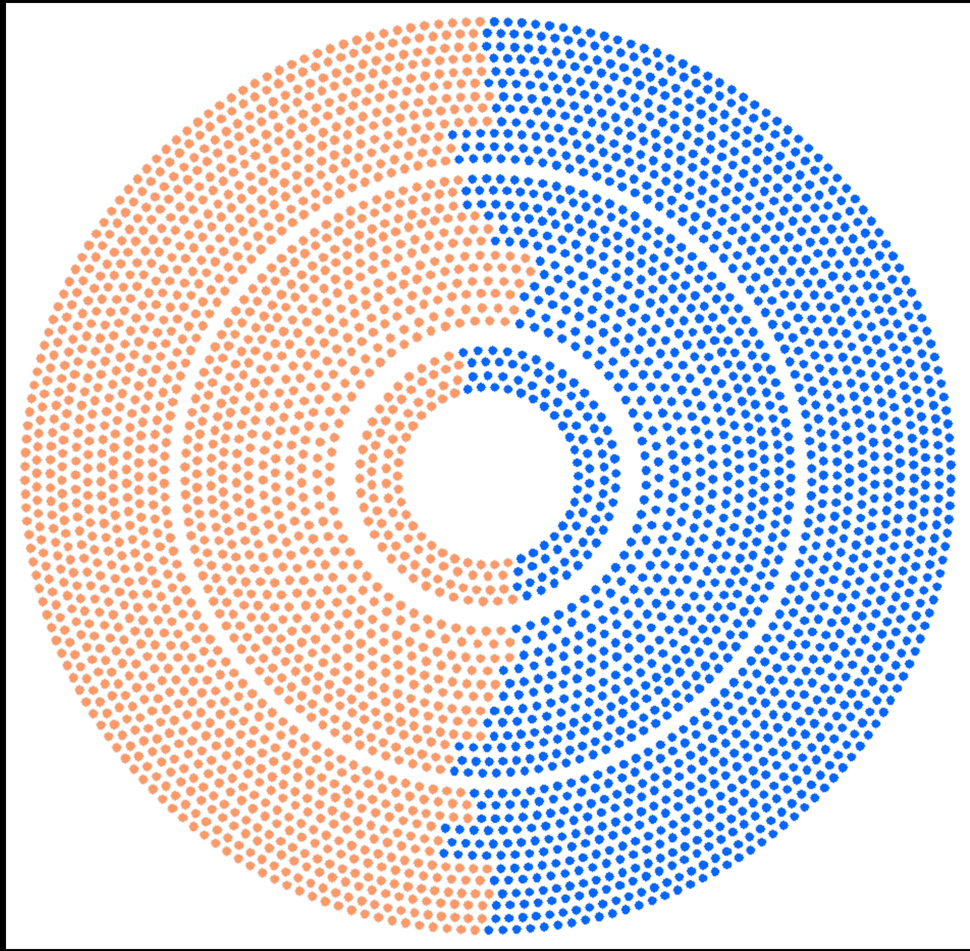




Stabilizing Gain in the Central Drift Chamber

- **Peak heights from Gaussian Process side** of the CDC show dramatic reduction in pressure dependence compared to **constant HV**

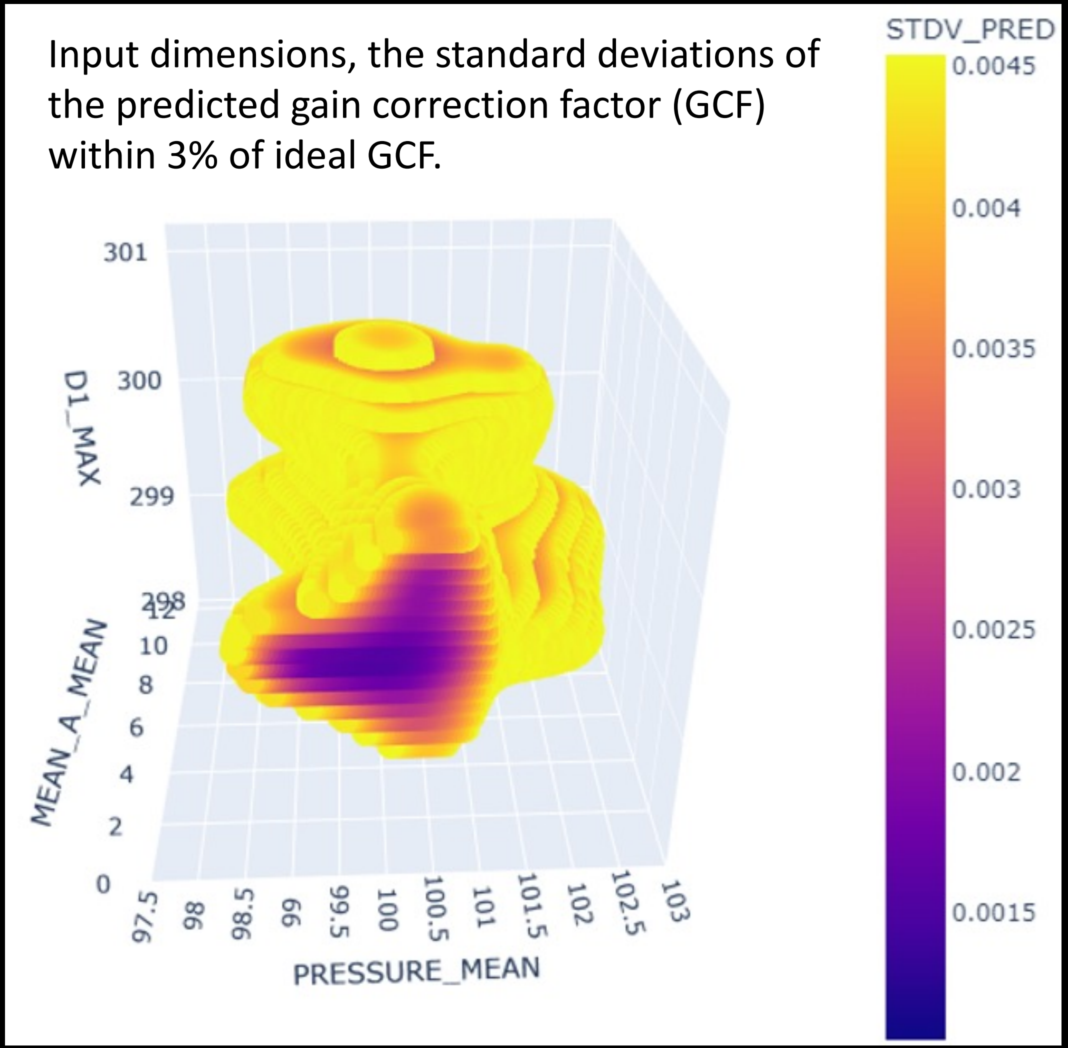




Schematic of downstream view of CDC, with straws HV control status indicated.

Considering UQ for control decisions

- We don’t not want to adjust high voltage to an "uncertain" value
- Gaussian Process provides uncertainty quantification
- Only apply a new calibration if the uncertainty is within the 3% of ideal gain correction factor otherwise, we extract to the closest prediction within tolerance
- This method will be implemented in the upcoming GlueX CPP run

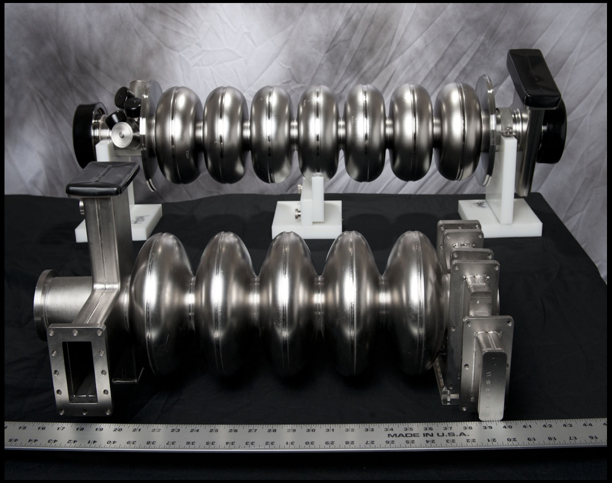


Input dimensions, the standard deviations of the predicted gain correction factor (GCF) within 3% of ideal GCF.

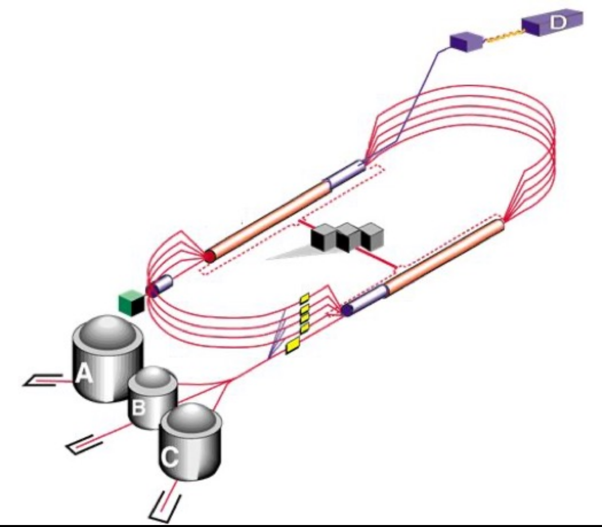
Multi-objective optimization for CEBAF

- SRF Cavities accelerate electrons via an RF standing wave
- Current method to set the cavity gradients
  - Fitting gradients vs fault rates on historical data
  - Heat generated by RF power input to the cavities is ignored

Suboptimal

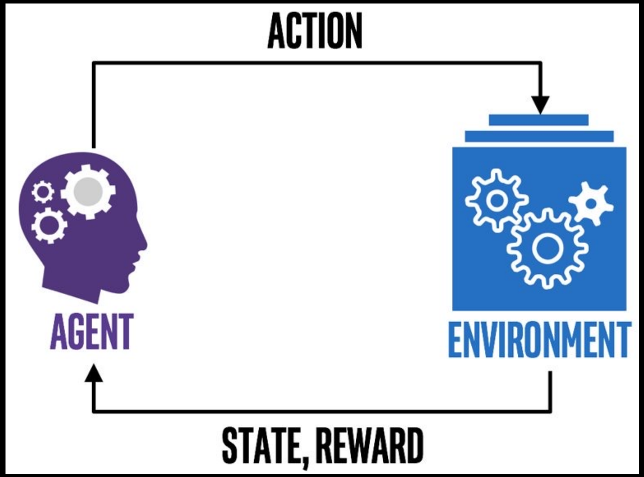


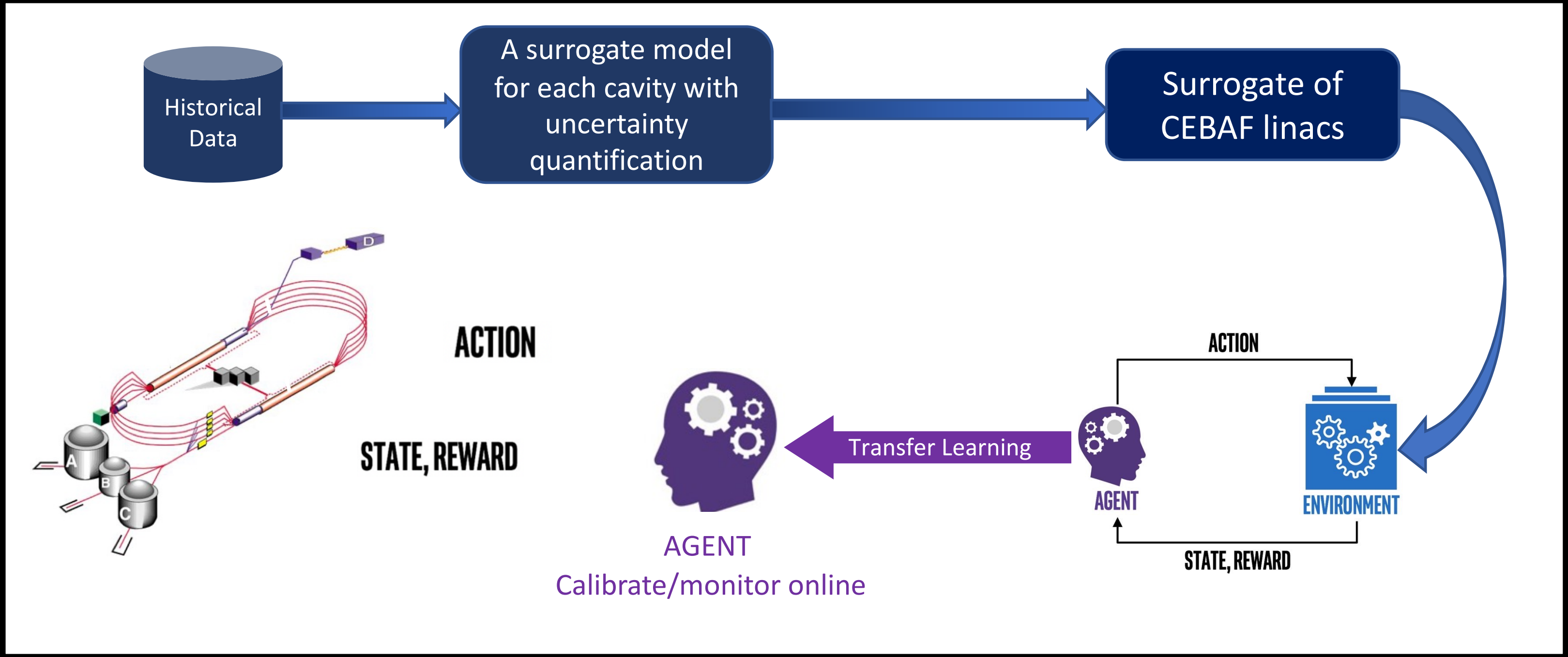
- Leads to higher heat load and trip rates.
- This problem is faced by all the SRF based accelerator facilities including JLab, SNS, FRIB, and LCLS II at SLAC, if the gradients are not optimal.



Multi-objective optimization for CEBAF

- Dynamic problems need dynamic solutions.
- Reinforcement Learning (RL) is a perfect fit and is proven to work well on optimization problems.





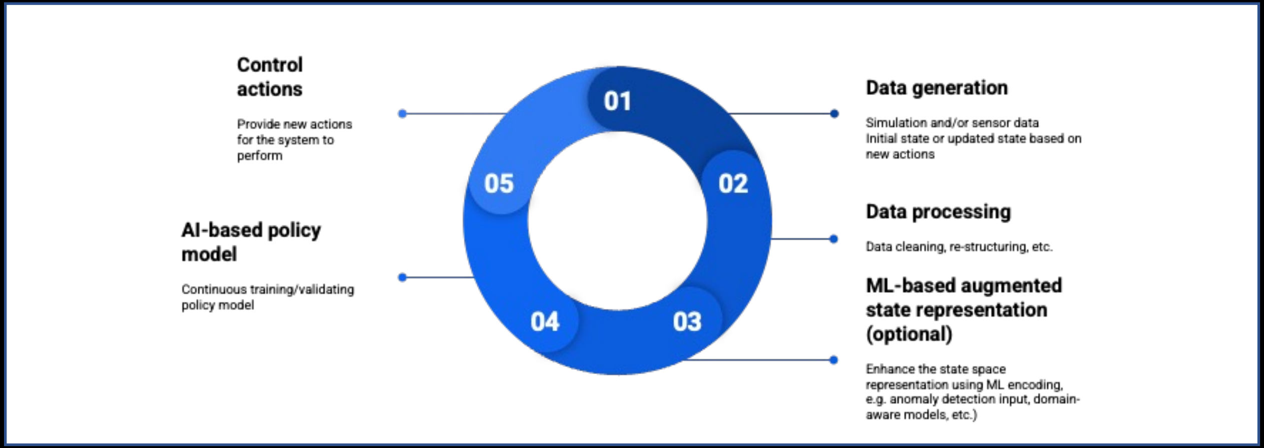
Critical components

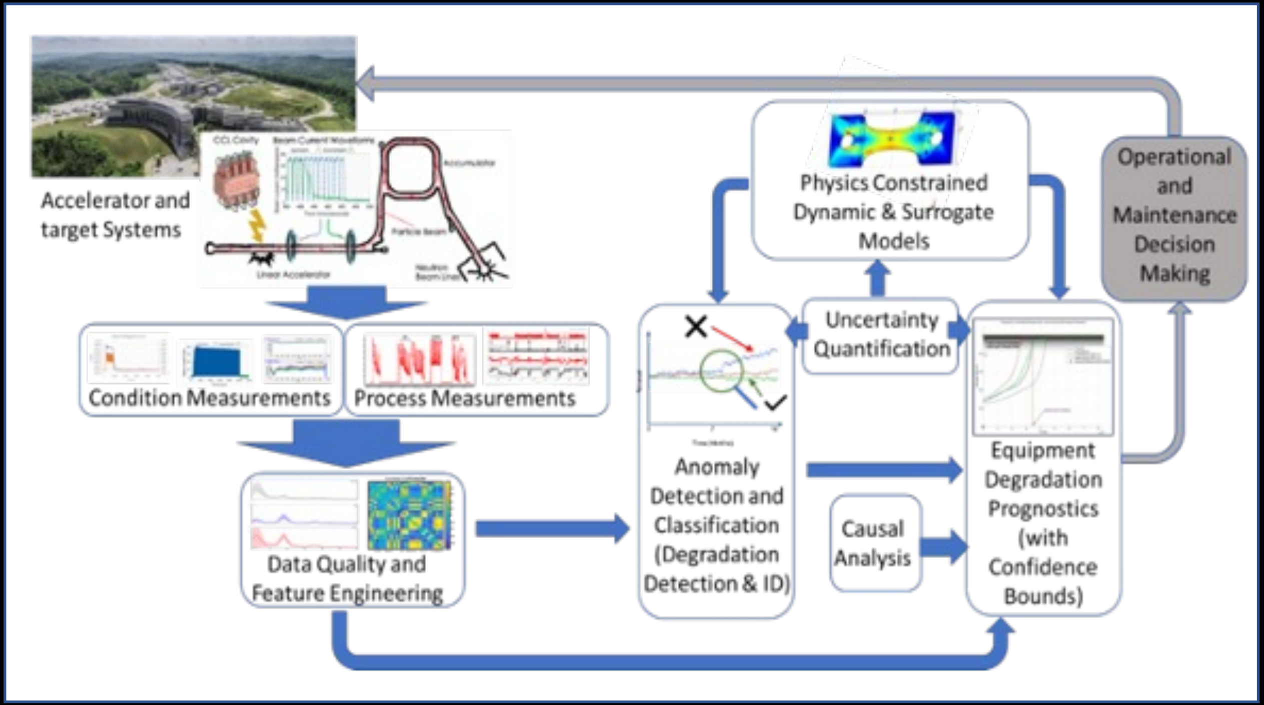
•**Multi objective optimization:** getting optimal set of gradients that minimizes both trips and heat

•**Q-modeling:** The ideas was to account for the Q0 dynamics with gradients within the surrogate models

Combining research efforts into a grander workflow

- Implement UQ methods at every level to understand what we don't understand
- Incorporate interpretability, explainability, and robustness to provide stable solutions
- Integrate into design & controls workflows to ensure optimization is within our knowlegde

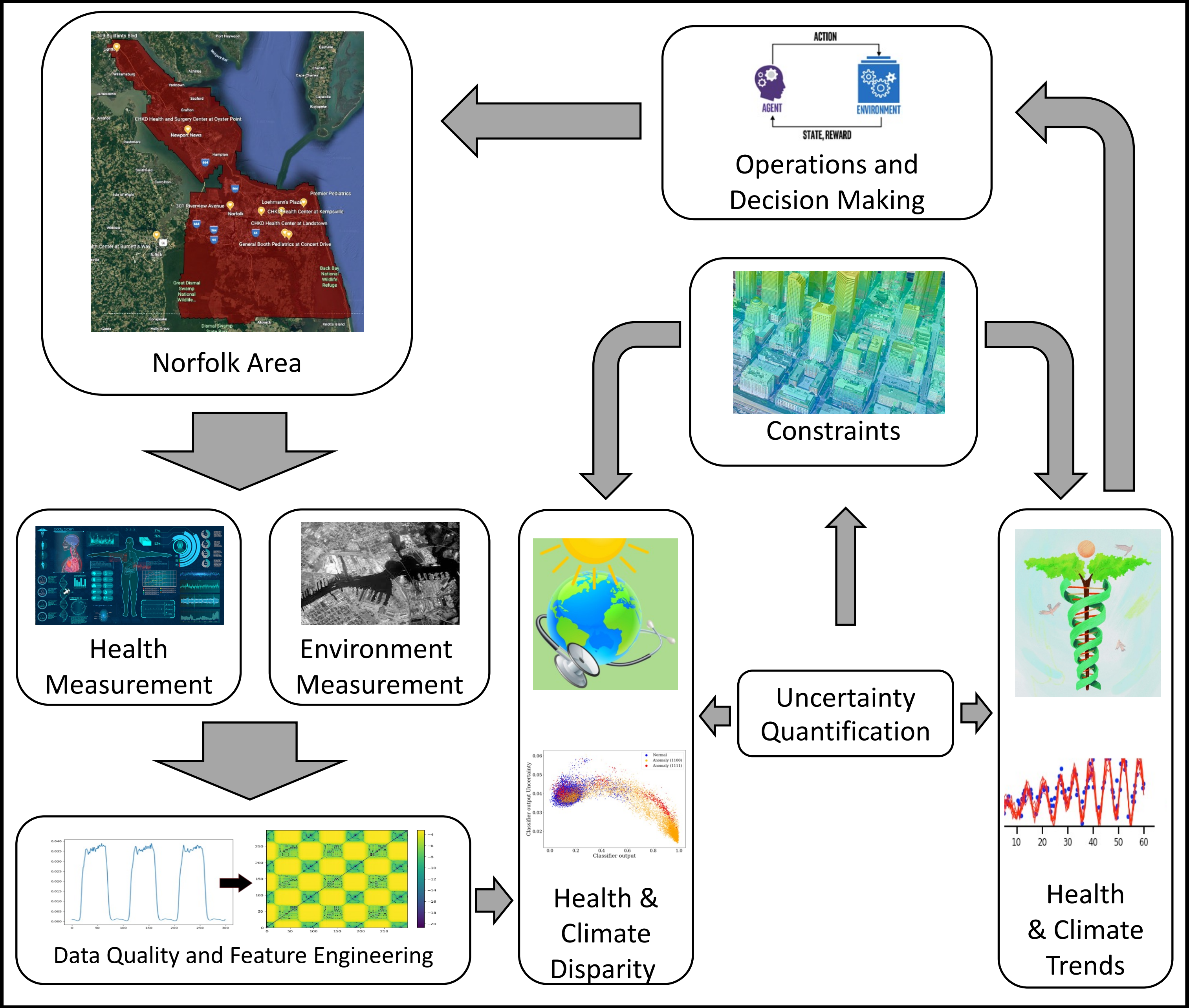




Applying data science workflow to Health and Climate Studies

- Collected multi-modal data. of the Norfolk area
- Develop multi-modal digital model of the Norfolk area
- Apply UQ and constraints methods

- Provide decision policy based on research topic



Final Remarks

- Cris Fanelli is the latest member of the team
  - New bridge position between W&M and Jlab
- Diana McSpadden
- Nikhil Kalra
- Two new postdocs
  - Yasir Alannazi
  - Abdullah Farhat

Hiring a new postdoc now

Contributed to several AI related workshops:

- JLab Experimental Hall Townhall
- JLab accelerator Townhall
- AI4EIC workshop
- AI@DOE workshop
- AI for Science and Security workshop
- AIRES
- Etc.

We are involved in several projects at JLab:

- Experimental Hall
- Theory
- Accelerator

We also have several projects outside of NP

We are always looking to collaborate.

Please contact us.

AI4EIC: Upcoming Event

AI4EIC https://eic.ai/events ai4eic.slack.com

ECCE AI WG ecceaiwg.slack.com

EIC Software: AI WG Meeting

Wednesday 22 Jun 2022, 09:00 → 11:05 US/Eastern

https://indico.bnl.gov/event/16073/

Description This meeting is topic-oriented and focused on uncertainty quantification.

We will use Zoom for the remote meeting: https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09

Meeting ID: 1614875218 Password: 925723

We will use live notes where everyone can edit comments/questions: https://docs.google.com/document/d/1iPNVJUuTjydop450nsHK78VuzilSzwB177KaXTNPips/edit

09:00	→ 09:05	**Introduction** **Speakers**: Cristiano Fanelli (MIT), Tanja Horn (Cath)	5m
09:05	→ 09:30	**Uncertainty aware ML-based models for accelerator studies (Anomaly detection and surrogate model for RL)** **Speaker**: Malachi Schram (Thomas Jefferson National Accelerator Facility)	25m
09:30	→ 09:40	Q/A	10m
09:40	→ 10:05	**Uncertainty Quantification for Machine Learning Applied to Data Analysis** **Speaker**: Benjamin Nachman (Lawrence Berkeley National Laboratory)	25m
10:05	→ 10:15	Q/A	10m
10:15	→ 10:40	**Inverse problems in nuclear tomography** **Speaker**: Nobuo Sato (Jefferson Lab)	25m
10:40	→ 10:50	Q/A	10m
10:50	→ 11:00	**Final Discussion**	10m

AI4EIC - October 10-14, 2022

2nd General Workshop on Artificial Intelligence for the Electron Ion Collider

Venue: William and Mary

Contacts:

support@eic.ai