

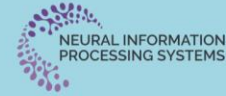
Uncertainty Guided Online Ensemble for Non-Stationary data Distributions in Fusion Science

Kishansingh Rajput^{1, 2}, Malachi Schram¹, Brian Sammulu³, Sen Lin²

¹Data Science Department, Thomas Jefferson National Accelerator Facility, Newport News, VA

²Department of Computer Science, University of Houston, Houston, TX

³General Atomics, San Diego, California, USA



Reliable uncertainty based guidance for online learners significantly enhances drift handling.

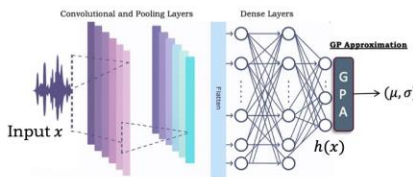
Motivation

- Need to handle data drifts in Fusion experiments
- Data drifts invalidate static ML models
- Online learning is critical for data driven Digital Twins
- Uncertainty Quantification is essential on OOD data

Use Case

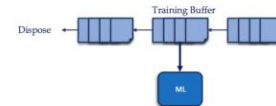
- Fusion shots can displace the outer structure (B-coils)
- Important to know expected deflection before running shots
- Using data from DIII-D National Fusion Facility
- Uncertainty aware model via Gaussian Process (GP) Approximation

Uncertainty Aware Modeling and Online Learning



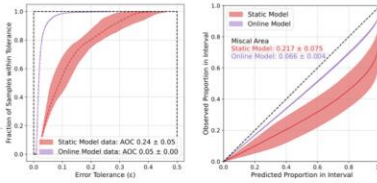
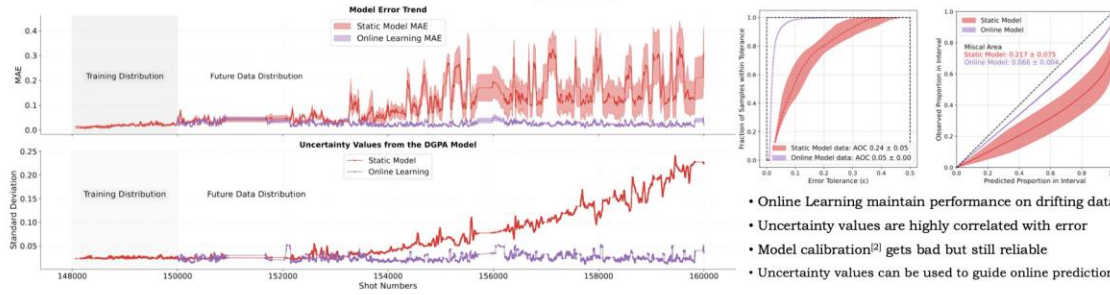
- Traditional GP scales $O(n^3)$
- Neural Networks are better suited to learn from large amount of data
- GP Approximation using a fixed sized Random Fourier Features kernel of size m
- GPA scales with $O(m^2)$ where $m \ll n$ and m is fixed
- Bi-Lipschitz constraint on latent representation (h_2) to preserve original input distances before feeding to GPA

$$L_1 \cdot |x_1 - x_2| \leq |h(x_1) - h(x_2)| \leq L_2 \cdot |x_1 - x_2|$$



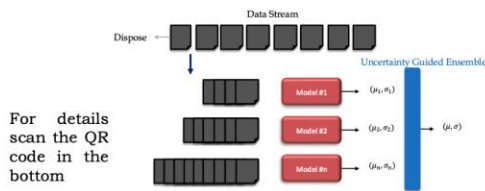
- Online fine-tuning using recent data buffer
- Forgetting on previous data is acceptable
- Fine-tuning need to be quick

Results

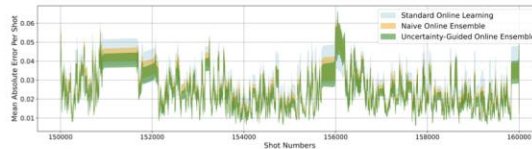


- Online Learning maintain performance on drifting data
- Uncertainty values are highly correlated with error
- Model calibration^[2] gets bad but still reliable
- Uncertainty values can be used to guide online predictions

Uncertainty Guided Online Ensemble



For details scan the QR code in the bottom



Metric	Static Model	Online Learning	Online Ensemble	UQ Guided OE
MAE ($\times 10^{-2}$)	11.88 \pm 2.50	2.41 \pm 0.05 (baseline)	2.26 \pm 0.02 (5.92%)	2.16 \pm 0.02 (10.24%)
MSE ($\times 10^{-2}$)	2.57 \pm 1.11	0.14 \pm 0.00 (baseline)	0.13 \pm 0.00 (8.66%)	0.12 \pm 0.00 (12.56%)
MAPE	22.44 \pm 4.61	4.18 \pm 0.10 (baseline)	3.92 \pm 0.05 (6.30%)	3.80 \pm 0.04 (9.08%)

Summary

- Online learning is critical to maintain performance of data-driven ML
- Uncertainty guided online ensemble outperforms standard online learning
- In future, we would like to deploy the model and evaluate online

References

1. Rajput, K., Schram, M., Sammulu, B., and Lin, S. Uncertainty guided online ensemble for non-stationary data streams in fusion science, 2025, url: <https://arxiv.org/abs/2511.0292>
2. Uncertainty Toolbox: an Open-Source Library for Assessing, Visualizing, and Improving Uncertainty Quantification, 2021, url: <https://arxiv.org/abs/2109.10254>

Acknowledgement

Work supported by Jefferson Lab and General Atomics internal research and development. This manuscript has been authored by Jefferson Science Associates (JSA) operating the Thomas Jefferson National Accelerator Facility for the U.S. Department of Energy under Contract No. DE-AC05-06OR223177. Additionally, this material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Award(s) DE-FC02-04ER54698, along with Office of Fusion Energy Sciences Award No. DE-SC0024426 (Fusion Data Platform). The US government retains, and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

