Accelerating Disruption Database Studies with Semi-Supervised Learning

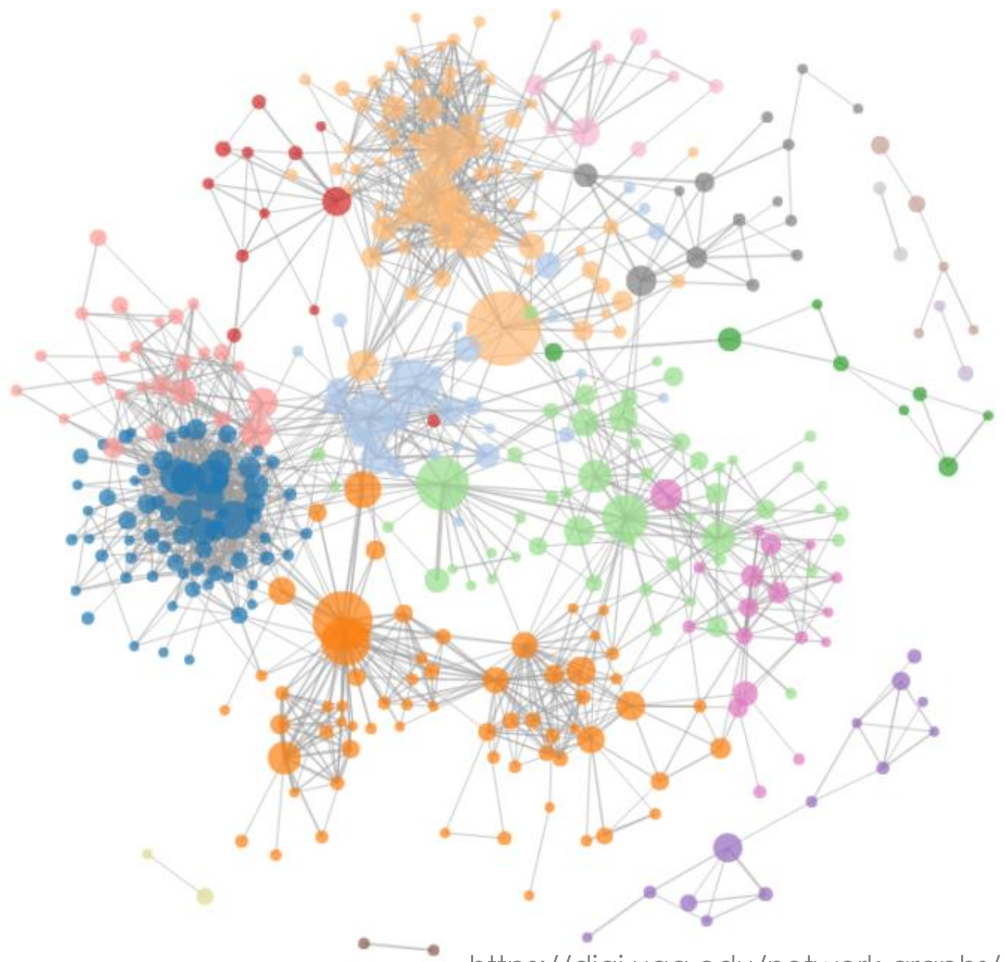
by

K.J. Montes[1*]

with J. Zhu[1], C. Rea[1], R.A. Tinguely[1], R. Sweeney[1], and R.S. Granetz[1]

[1]MIT Plasma Science and Fusion Center, USA

*kmontes@mit.edu



https://digi.uga.edu/network-graphs/

Presented at the

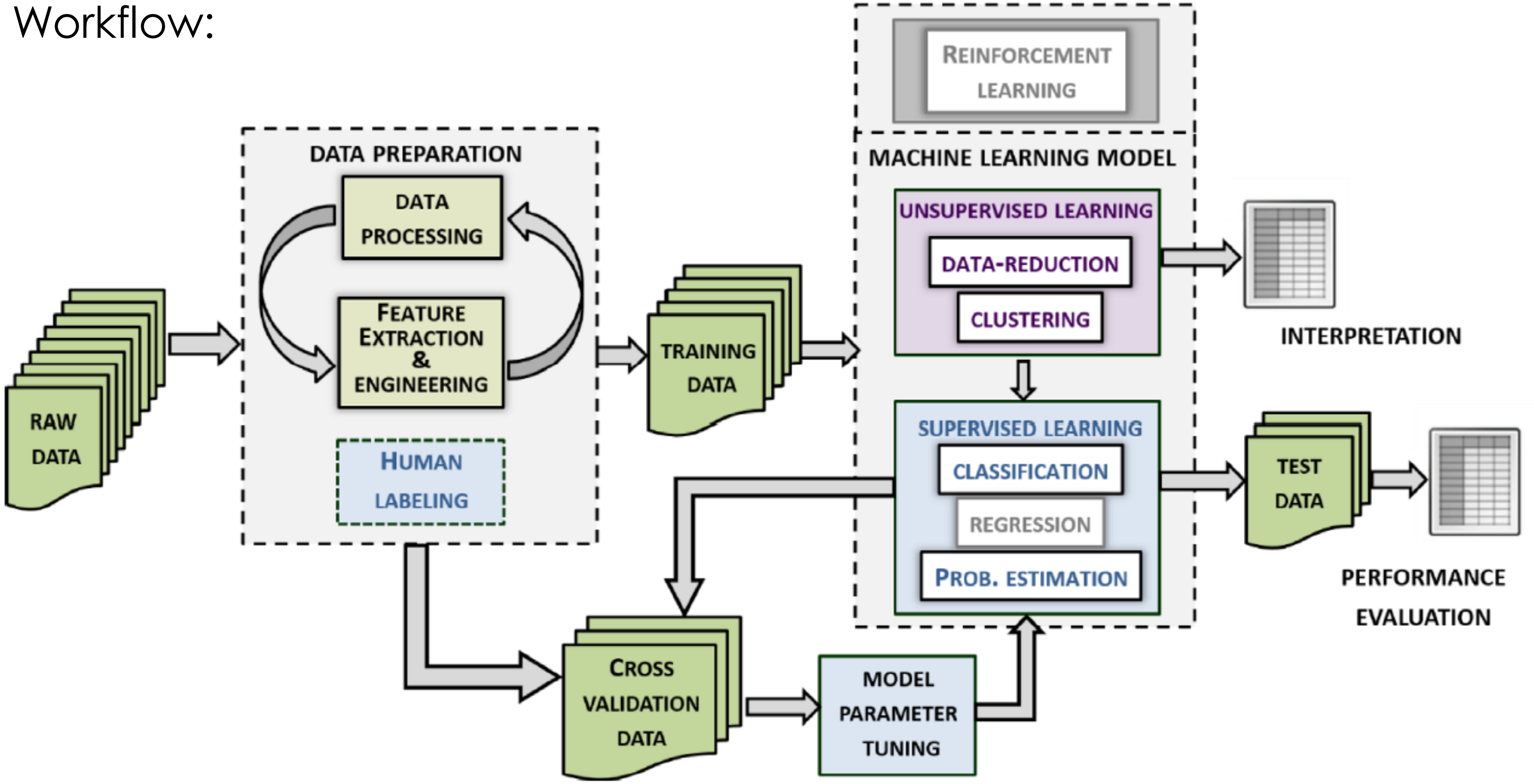
IAEA Technical Meeting on Plasma Disruptions and their Mitigation

July 20-23, 2020

Work supported by U.S. DOE under DE-FC02-04ER54698.

Data-driven disruption prediction requires large labeled databases

Typical Workflow:

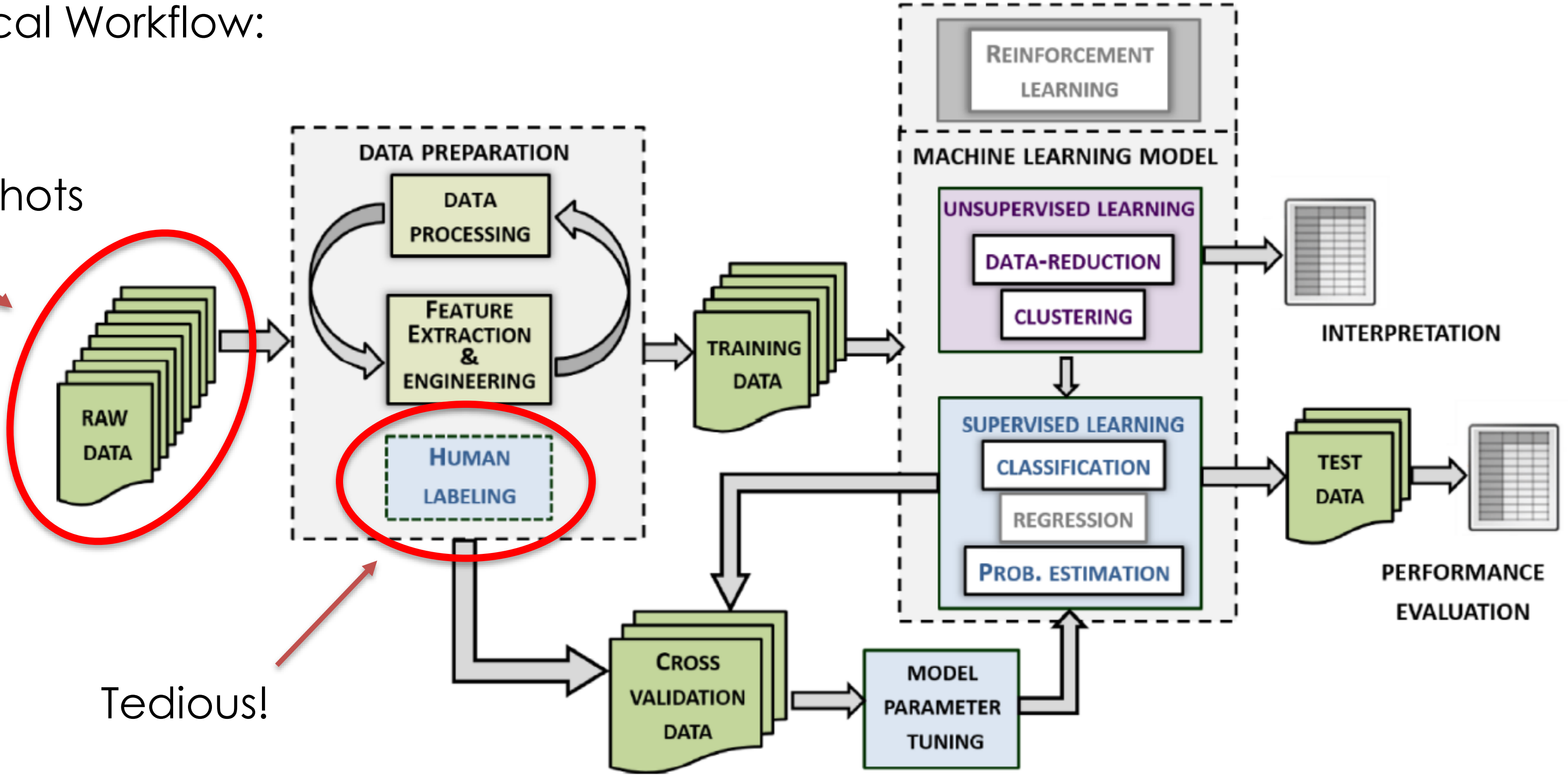


A. Pau *et al* 2019 *Nucl. Fusion* **59** 106017 (doi)

Data-driven disruption prediction requires large labeled databases

Typical Workflow:

$\sim 10^3 - 10^4$ shots

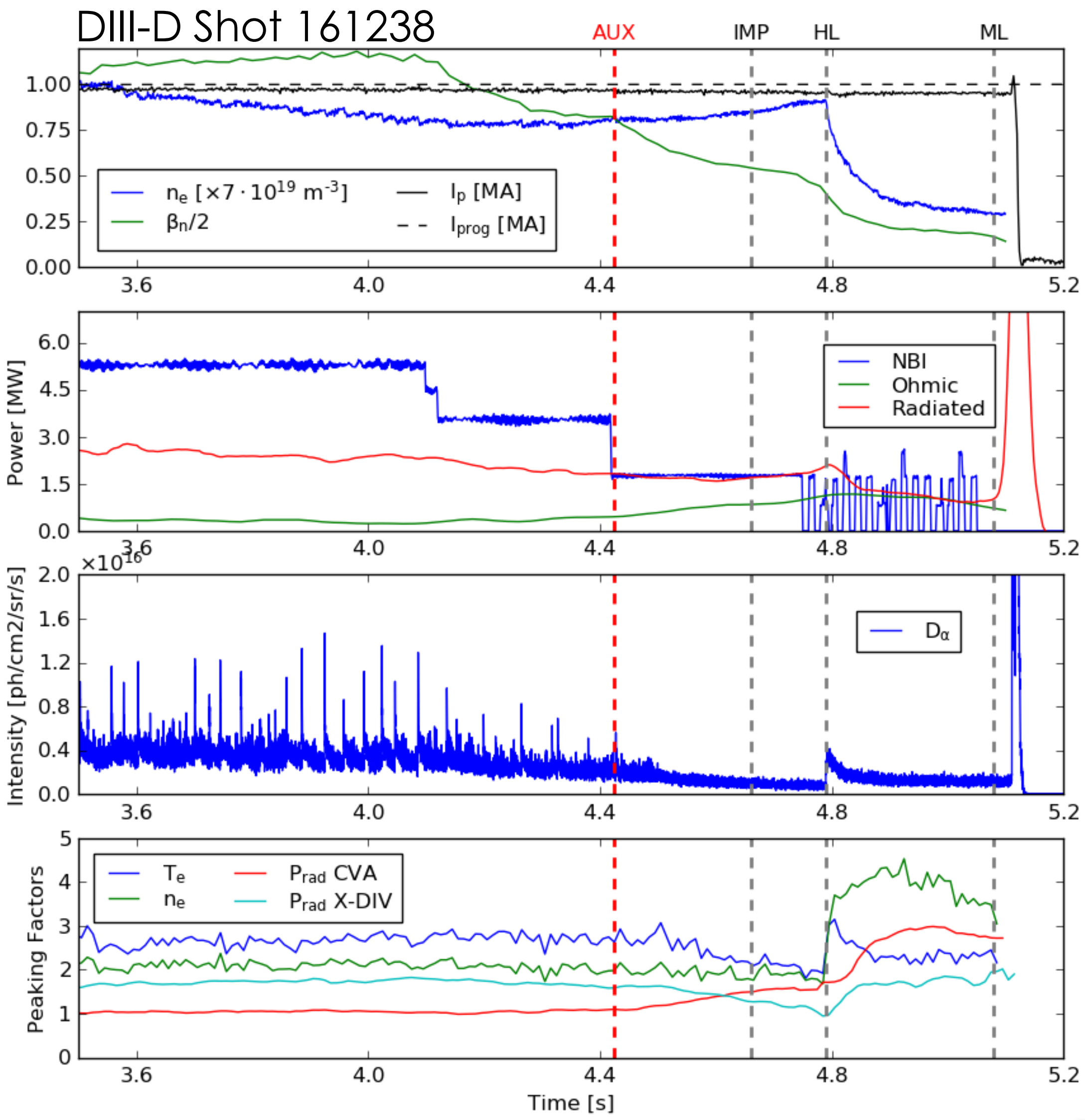


Tedious!

A. Pau *et al* 2019 *Nucl. Fusion* **59** 106017 (doi)

Labeled disruption precursor event chains can further understanding

- **Built dataset of manually labeled disruption precursors**
  - ~ 300 discharges from DIII-D 2015 & 2016
  - Recorded start time and type of each event
- **Inspired by study of disruption causes on JET[1] that labeled 2309 discharges!**
  - Later extended[2] to complement & interpret a machine-learning disruption predictor



[1] P.C. de Vries *et al* 2011 *Nucl. Fusion* **51** 053018 (doi)
[2] A. Pau *et al* 2019 *Nucl. Fusion* **59** 106017 (doi)

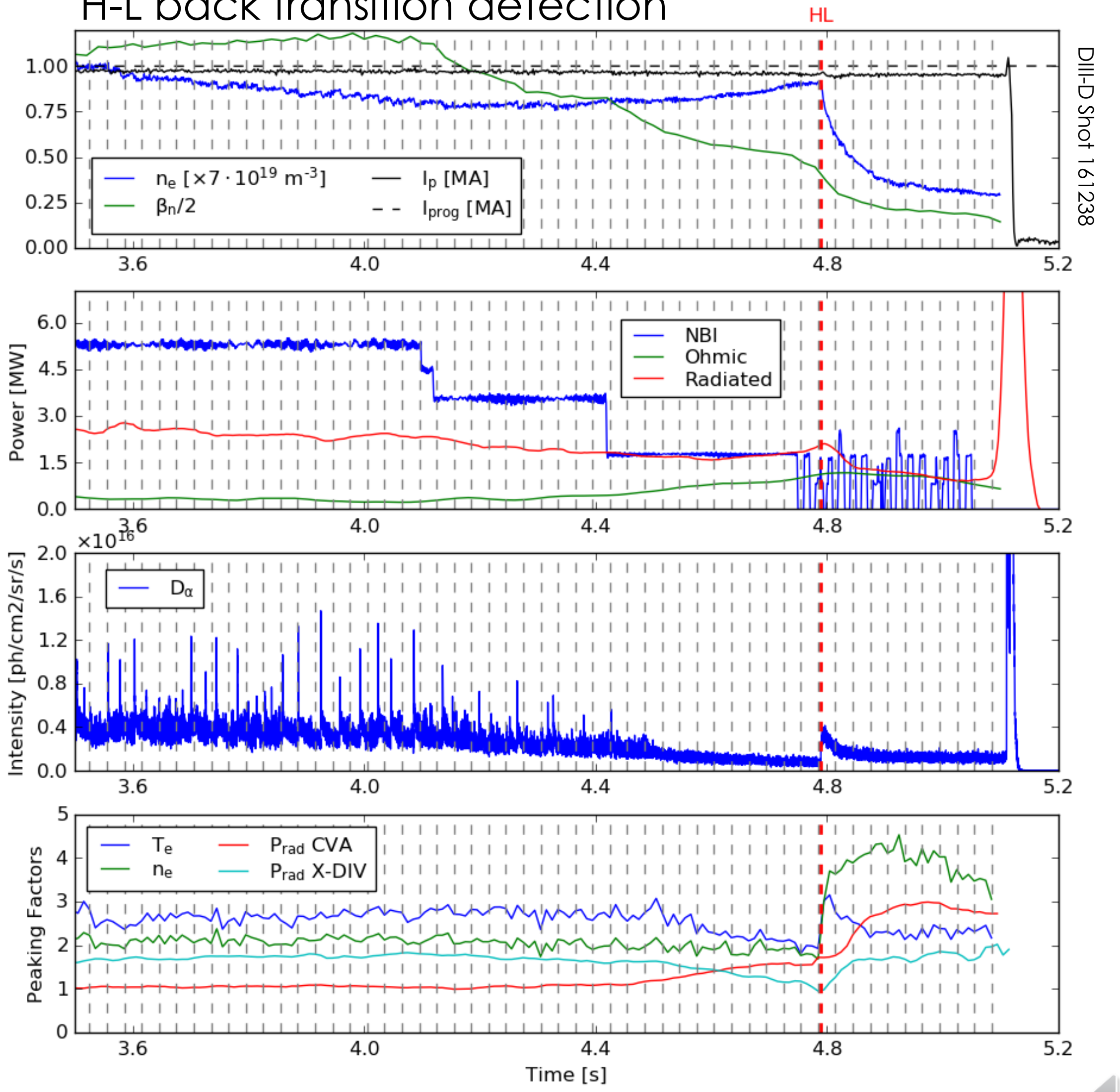
Semi-supervised learning on time-sequences helps detect/label events

- **Sample time sequences from each shot (endpoints are shown) with…**
  - Duration > event timescale
  - # of steps > event resolution
- **Choose N signals to identify event. Each sequence $\vec{x}_i$ now a point in high-D space:**

$$\vec{x}_i \in \mathbb{R}^{N\cdot(\#\ of\ steps)}$$

- **Dataset was <u>standardized</u> (scaled & offset so that $\mu = 0,\ \sigma = 1$)**

H-L back transition detection



Semi-supervised learning on time-sequences helps detect/label events

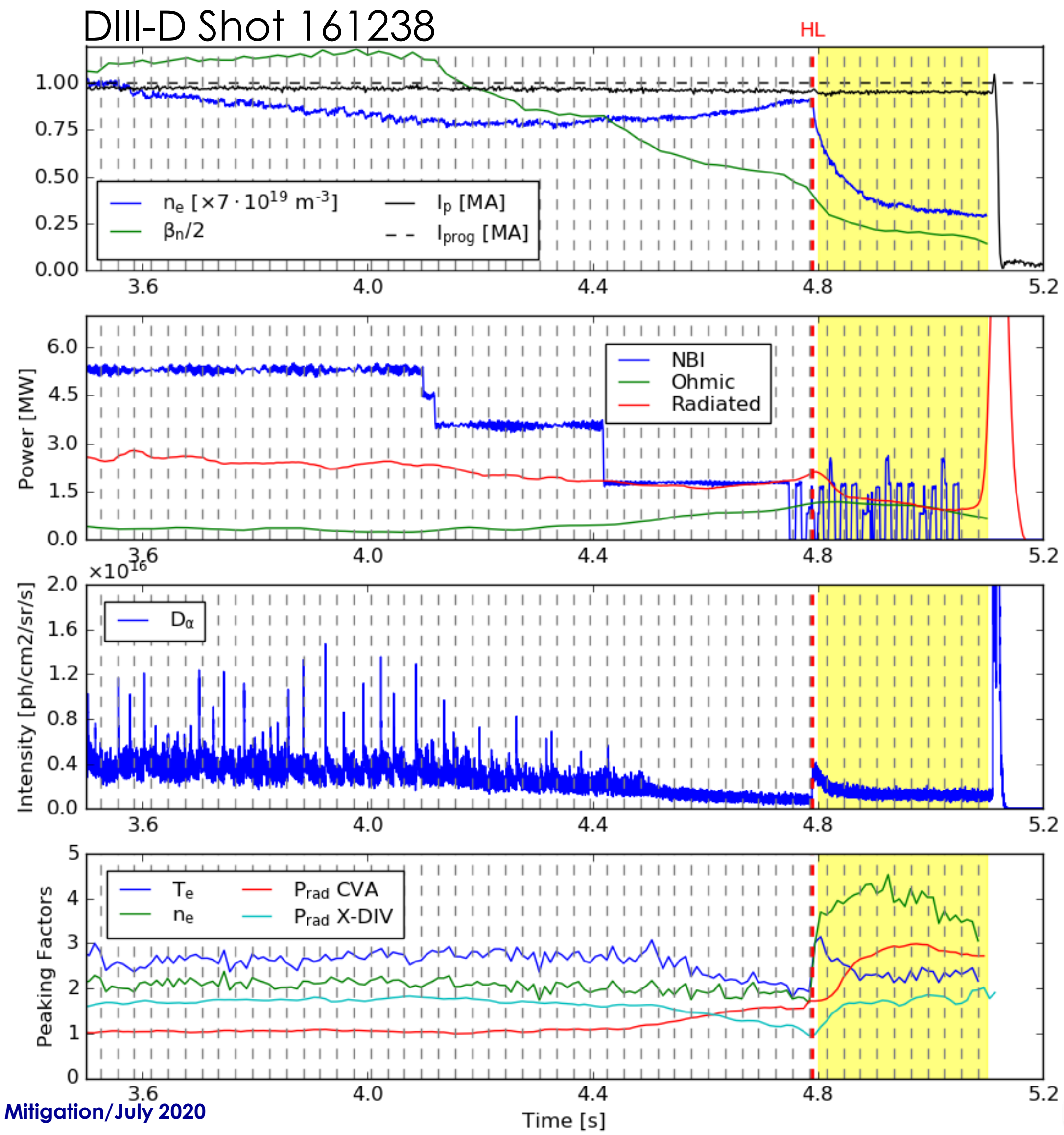
- **Assign ℓ sequences from this shot a class:**
  - Positive ($y_i = 1$) if it overlaps with event
  - Negative ($y_i = -1$) otherwise

$$Y_L = \{y_1, \cdots, y_\ell\}$$

- **Assign a placeholder class to u sequences from unlabeled shots**
  - Unobserved ($y_i = 0$)

$$Y_U = \{y_{\ell+1}, \cdots, y_{\ell+u}\}$$

- **<u>Goal of Semi-Supervised Learning</u>:**
  - Infer Y_U using all $\vec{x}_i \in X$ and Y_L, typically for cases where most data unlabeled $(\ell \ll u)$



DIII-D NATIONAL FUSION FACILITY

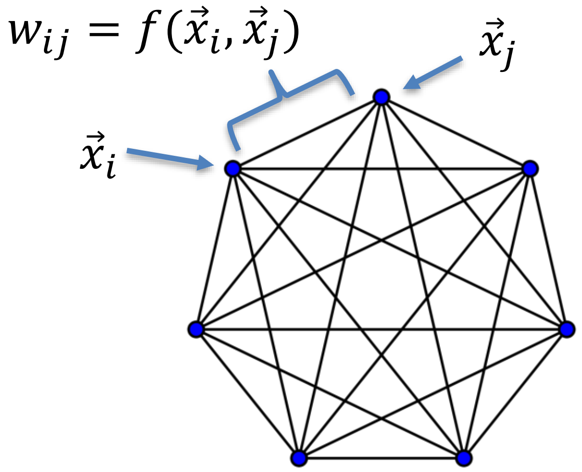
Label Propagation Algorithm

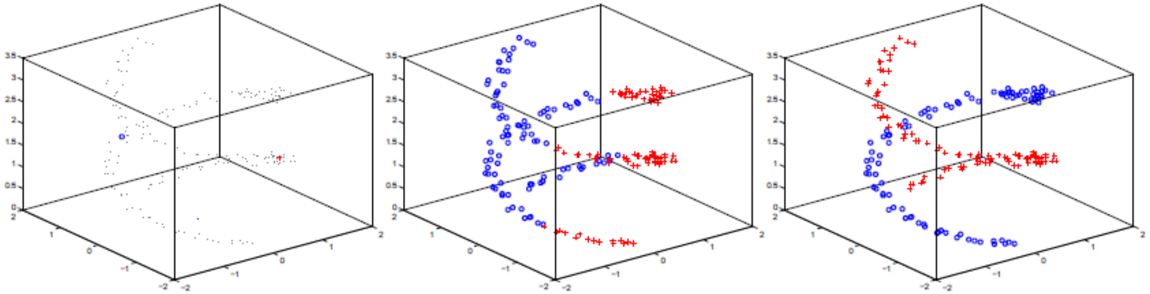
- **<u>Key Assumption</u>: data points that lie close together have similar labels**

1. **Visualize dataset as fully connected graph**
   - Nodes are data points with values $0 \leq Y_i \leq 1$ representing probability $\vec{x}_i$ is in positive class
   - Edges are weighted by Euclidean proximity
2. **Algorithm iteratively updates Y with transition matrix T**

$$T_{ij} = P(j \rightarrow i) \propto w_{ij}$$

3. **On each iteration, reset (clamp) originally labeled data $Y_1, \ldots, Y_\ell$ to original value**



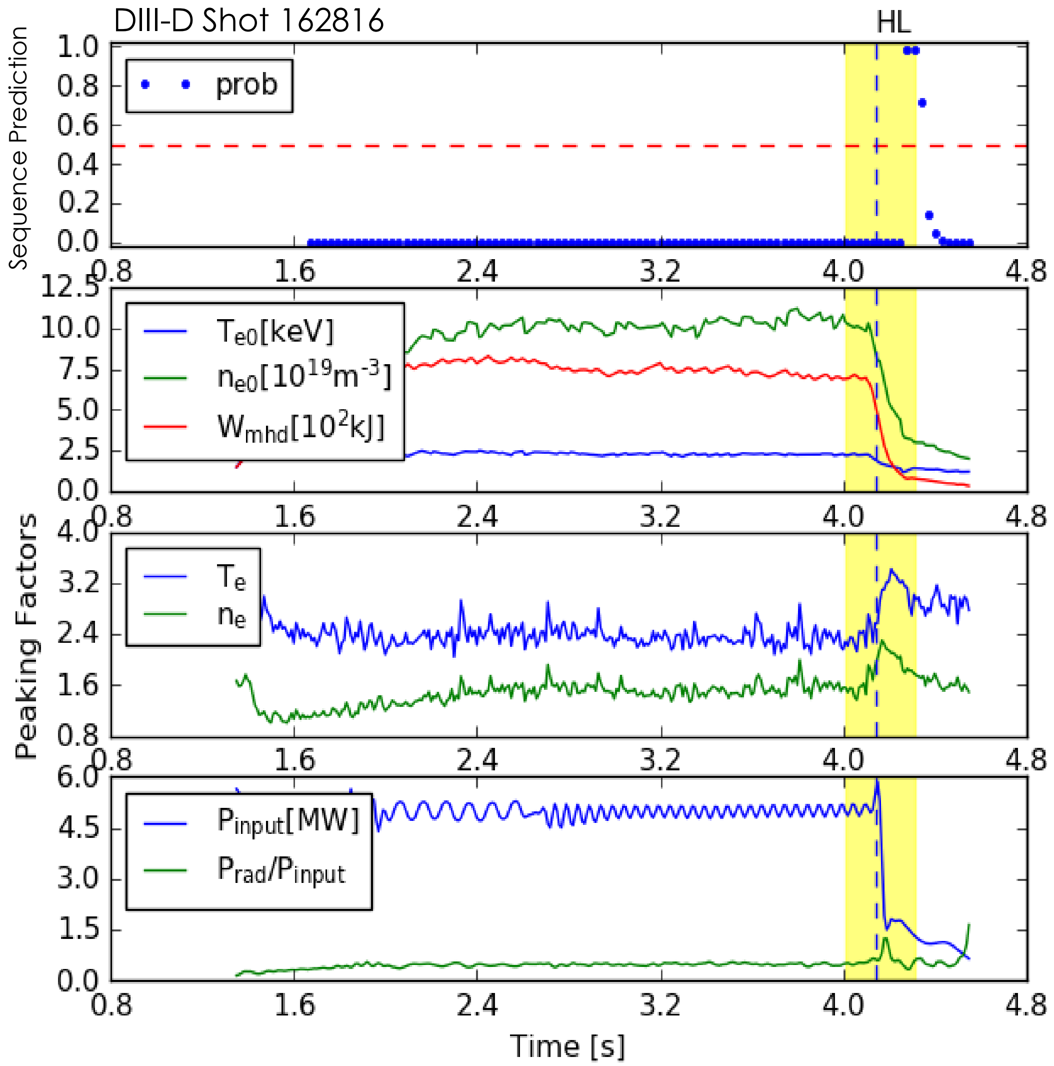


(a) The data (b) kNN (c) Label propagation

[1] X. Zhu and Z. Ghahramani 2002 *Technical Report CMU-CALD-02-107* (doi)

Applied label spreading [1] to detect H-L back transitions

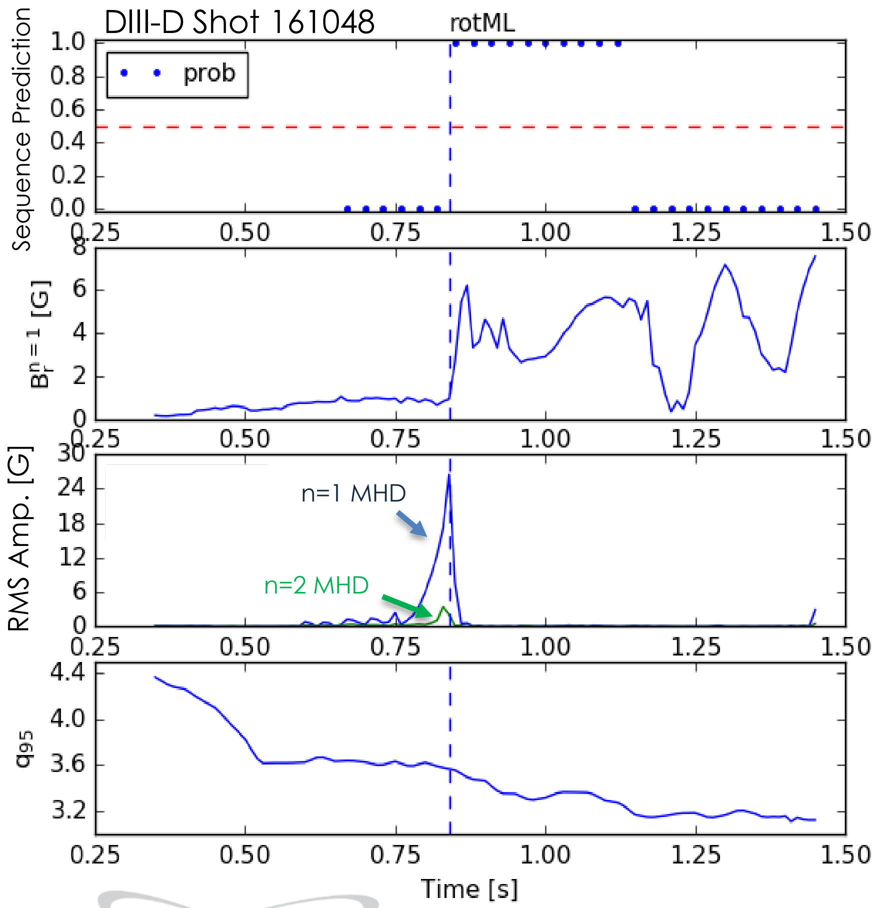
- **Event Prevalence:** $\sim 74\%\ (206/277)$ **of shots**
- **7 signals used, 6 time steps/sequence (42-D)**
- **Initially labeled 1.5% of shots**
  - Example shot 161238, along with 2 others with H-L transition & 1 without
- **Detection interval highlighted**
  - Remember, sequences depicted by endpoints
- $\sim 91\%$ **true positive rate (TPR)**
  - Fraction of shots w/ H-L back transition that had a successful detection
- $\sim 25\%$ **false positive rate (FPR)**
  - High-end estimate (for nuance, see extra slides)



[1] D. Zhou *et al* 2004 *Learning with local and global consistency* (doi)

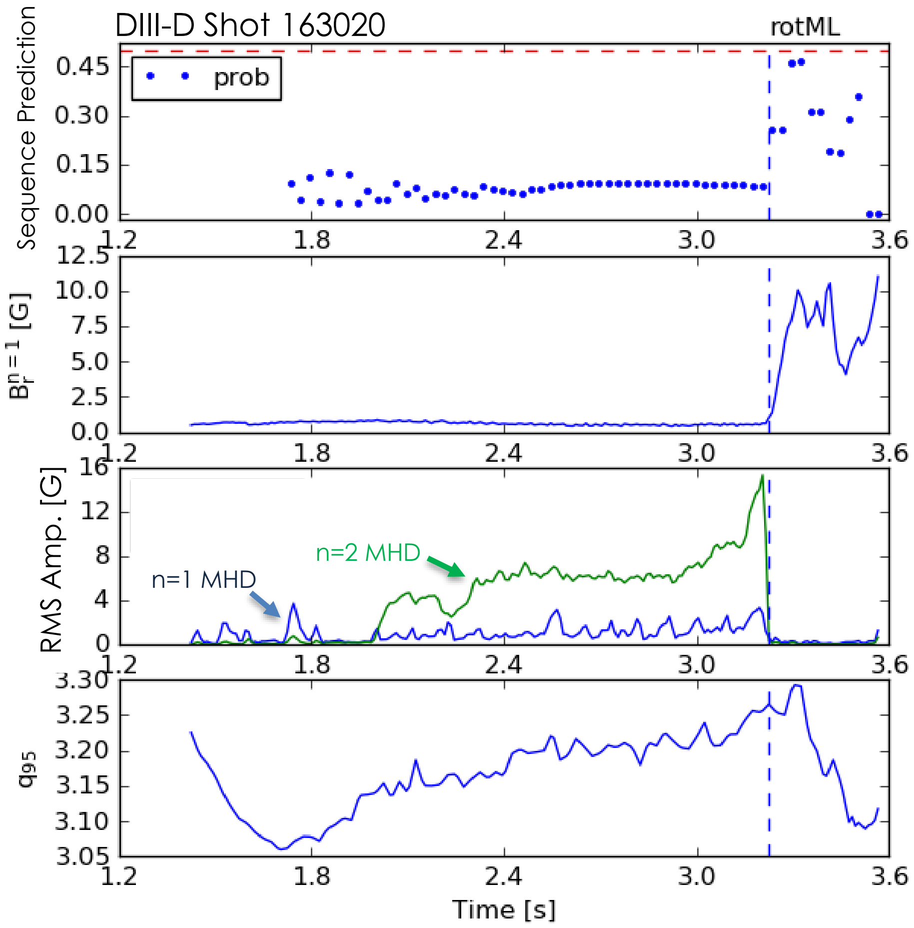
Applied same algorithm to detect locked modes with rotating precursors

- **Prevalence:** $\sim 65\%\ (180/275)$ **of shots**
- **Started with only 1 shot labeled (161048) and searched for marginal detections**



$$TPR = 86\%$$

$$FPR = 15\%$$



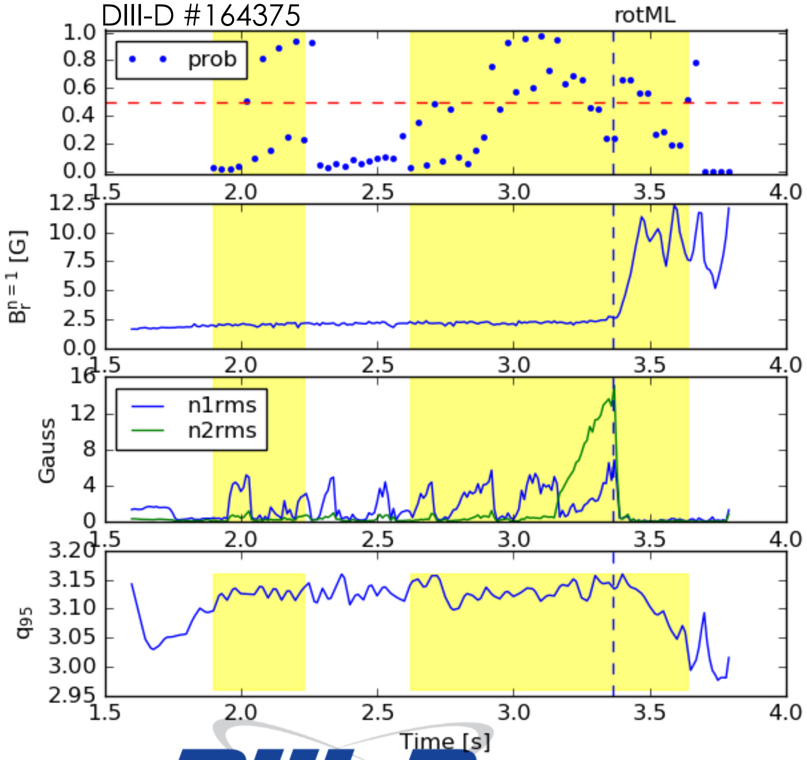
Iterative labeling method shows performance increase as shots added

- **Then, added the 'marginal' shot (163020) to the set of initial labels and retrained algorithm**

$TPR = 86\%$

$FPR = 15\%$

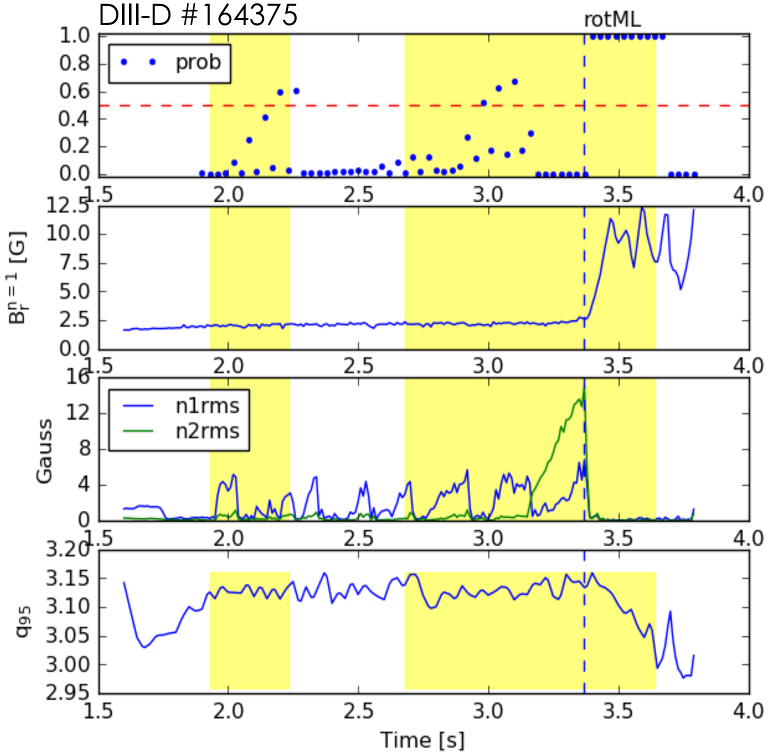
1st iteration



$TPR = 92\%$

$FPR = 16\%$

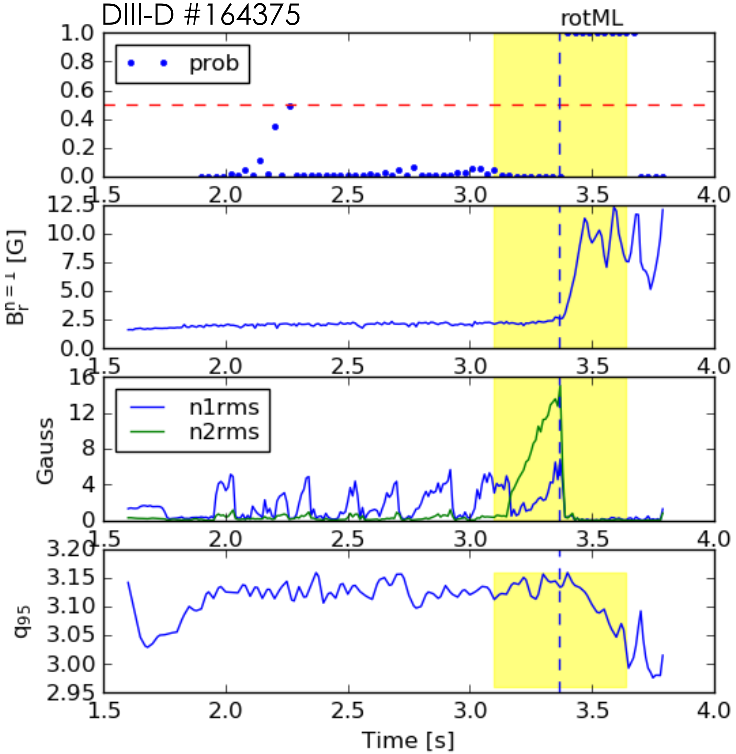
2nd iteration



$TPR = 94\%$

$FPR = 9\%$

9th iteration



DIII-D NATIONAL FUSION FACILITY

Application to core radiative collapses, a low prevalence event

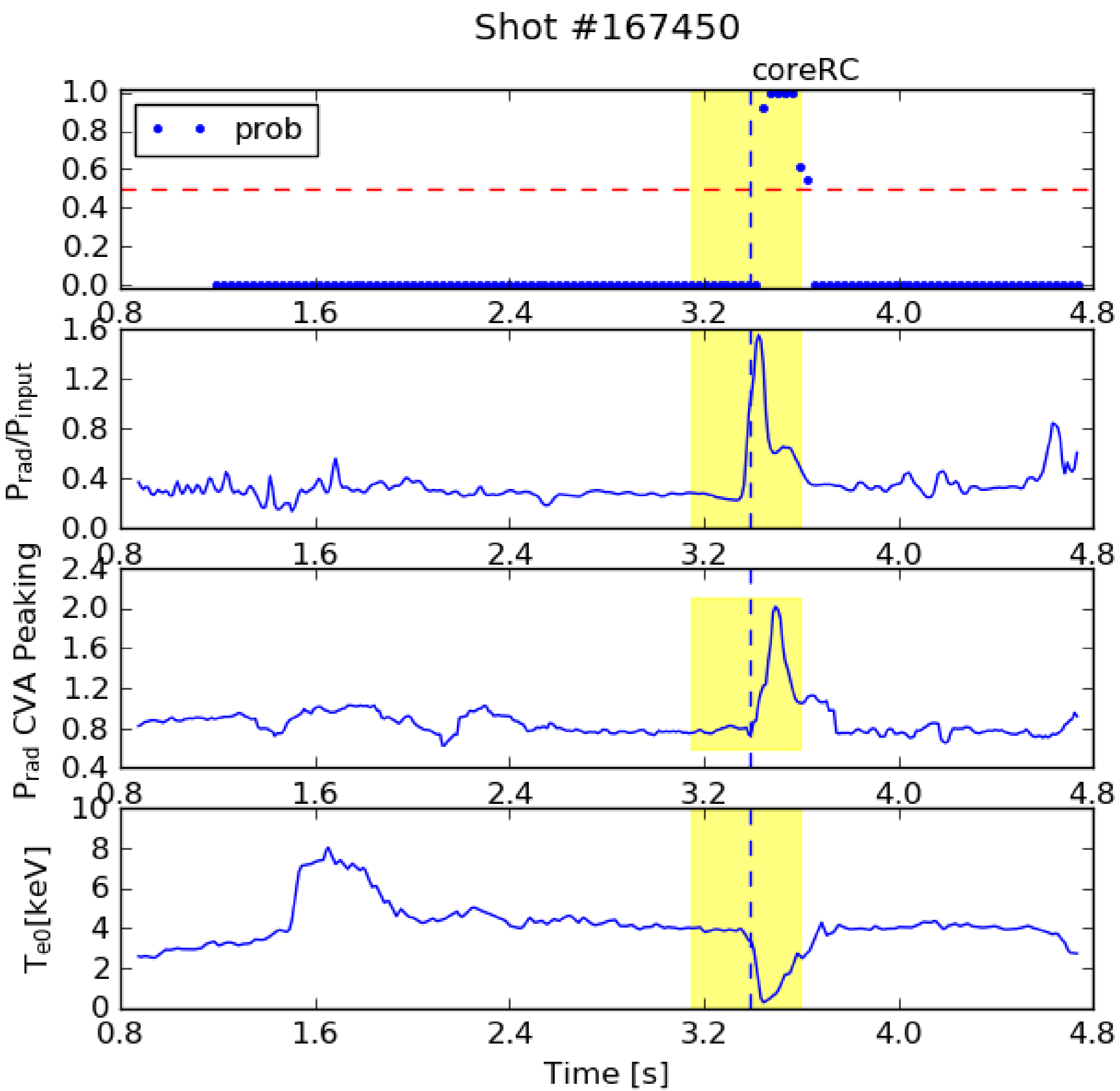
- **Prevalence: $\sim 8\% \ (25/294)$ of shots**
- **3 signals (18-D space)**
- **3 initial labels (2 with event, 1 without)**

$$TPR = 68\% \text{ (15/22 shots)}$$

$$FPR = 9\% \text{ (24/269 shots)}$$

- **Restricting dataset to shots w/ detections increases event prevalence by 5 times!**
  - Conservative estimate (see extra slides)

- **Can also be used as an experimental search engine (see extra slides)**



DIII-D NATIONAL FUSION FACILITY

Summary

- **Event databases can contribute to progress in disruption avoidance**
- **Reliable detections of <u>multiple events</u> demonstrated using the <u>same algorithm</u>**
  - Success with H-L transitions, initially rotating locked modes, and core radiative collapses
  - Suggests extension to arbitrary events is promising
- **Label spreading shows reasonable performance with little initial information**
  - All applications shown used only 1-5% of samples as initial labels
- **Performance increase observed as labeled examples are added**
  - "Pull up by the bootstraps" dataset construction method
- **Future work…**
  - Test other kernel functions and compare performance (see extra slides)
  - Finish developing OMFIT[1] module to share work & apply to arbitrary events
    - In progress (see extra slides for details)
  - Extend this analysis to a larger dataset to build an events database for avoidance studies

[1] O. Meneghini *et al* 2015 *Nucl. Fusion,* **55**, 083008 (2015) (doi)

Extra Slides

Label Spreading Algorithm – how is it different?

- **Uses a modified transition step with a modified clamping procedure**[1]
  - Given each edge weight w_{ij}, the transition matrix T has elements

$$T_{ij} = \frac{w_{ij}}{\left(\sum_k w_{ik}\right)^{1/2}\left(\sum_k w_{jk}\right)^{1/2}}$$

  - Let Y^* be the initial probability vector Y (either 1 for manually labeled, or 0 otherwise). At iteration t, update Y according to the rule:

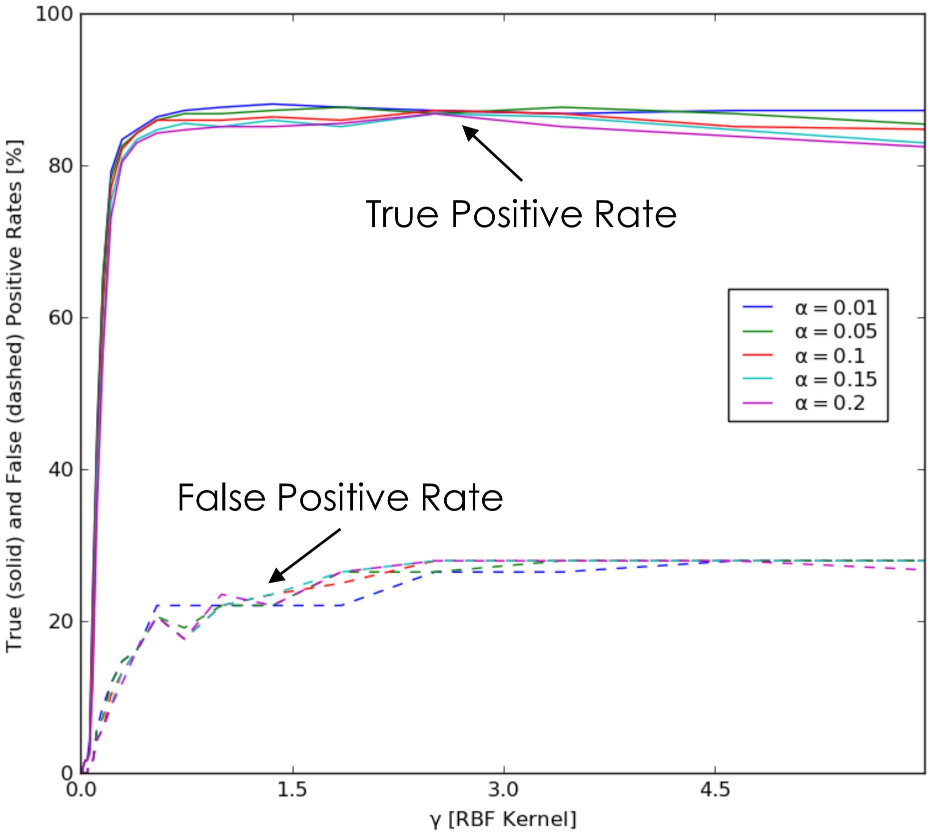
$$Y(t+1) = \alpha T Y(t) + (1-\alpha)Y^* \qquad [0 < \alpha < 1]$$

  - Here, α is called the **clamping factor**. Whereas label propagation performs a hard reset, or clamping, on each iteration, label spreading has a soft clamping effect.
    - α is chosen by the user. It can be changed to yield a softer clamping effect, allowing the algorithm to change the weight of the true ground labeled data to some degree[2]

[1] D. Zhou *et al* 2004 *Learning with local and global consistency* (doi)
[2] https://scikit-learn.org/stable/modules/label_propagation.html

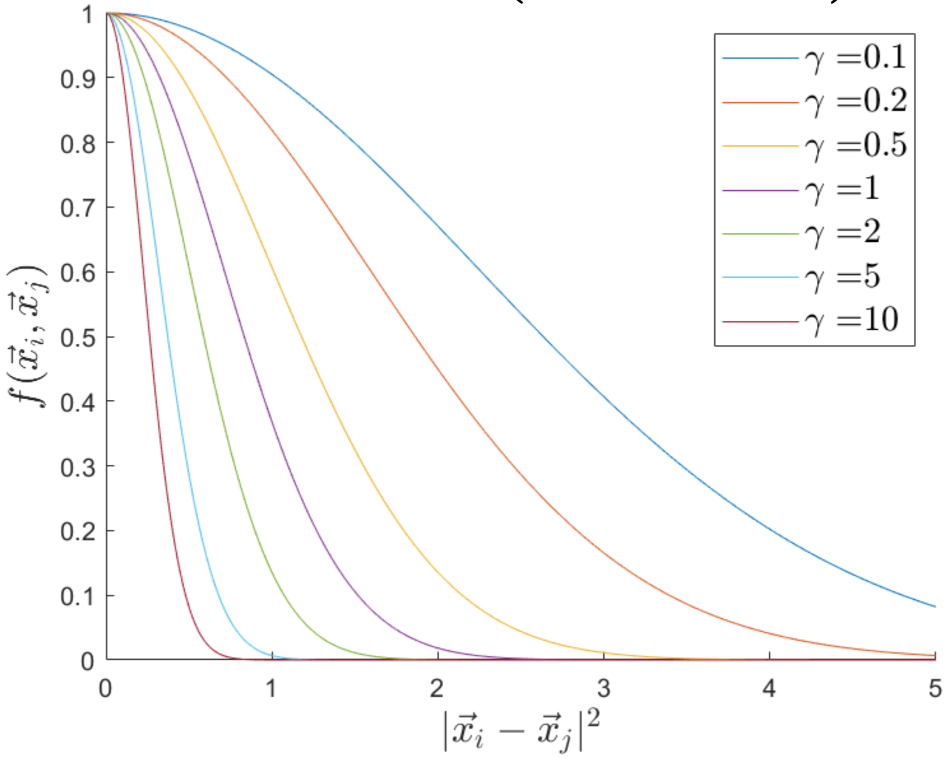
Performance is Sensitive to Kernel, Less Sensitive to Clamping Factor (α)

- **Performances increases with decreasing kernel width until hitting a plateau**
- **H-L performance shown, but same trend holds for other events**



Radial Basis Function

$$f(\vec{x}_i, \vec{x}_j) = \exp\left(-\gamma \left\|\vec{x}_i - \vec{x}_j\right\|^2\right)$$



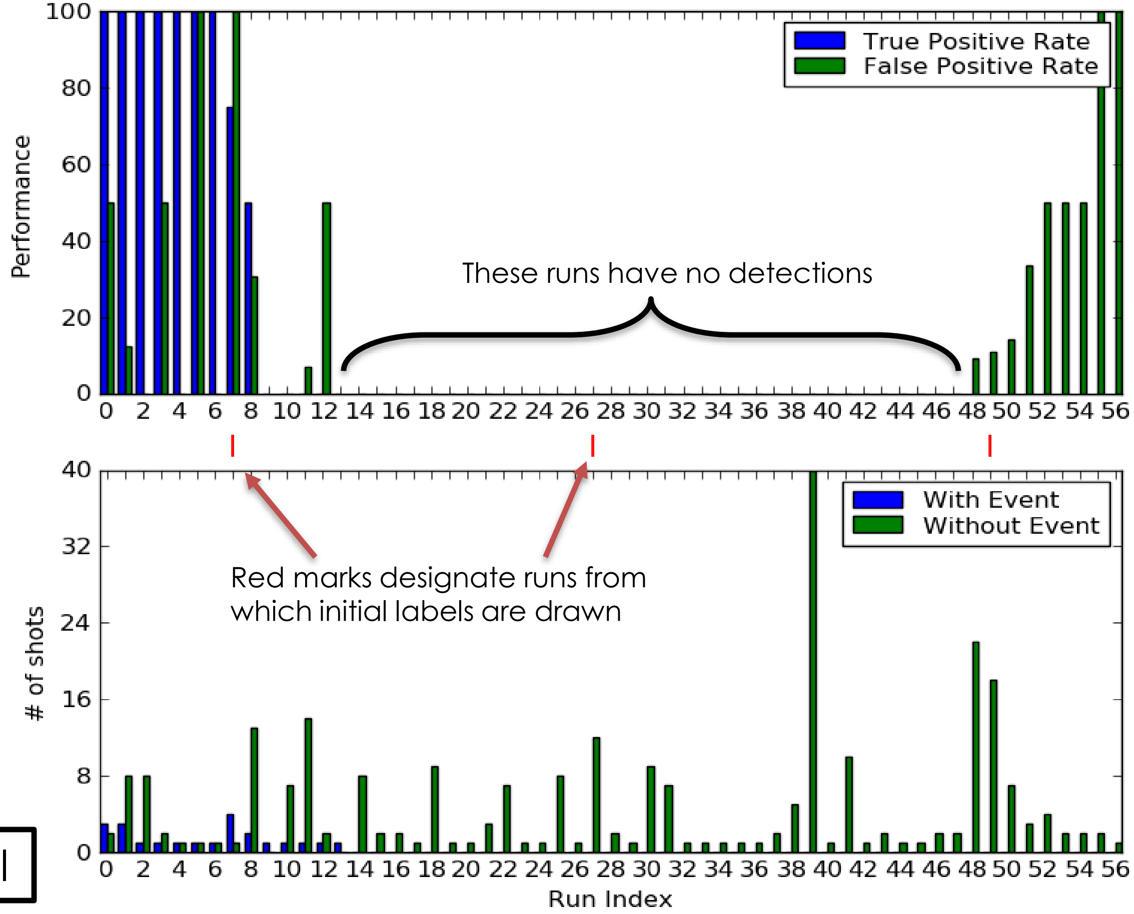
Can use event detection to find experiments with common themes & dynamics

- **57 runs from 2015/2016 campaigns with disruptions in core radiative collapse dataset**
  - Sorted by performance (best/worst, left/right)
- **Correlation b/w experiment & event found**
  - Impurity accumulation, high-Z divertor, ELM control associated w/ core radiative collapse

<u>Best Performing Run Days</u>

0 Isotope scaling L-H and H-L power thresholds
1 Impact of High Power AT Operation … using W Tiles
2 End of Metal tiles campaign tasks
3 High Frequency D2 Pellet ELM Pacing
4 Plasma Startup and Systems Checkout - Day 2
5 Detachment onset at the inner and outer divertor
6 Impurity Granule Injector and D2 Fueling
7 Zonal Flow Damping in L-H Transitions
8 Effect of RMP ELM Control on W Divertor Erosion…

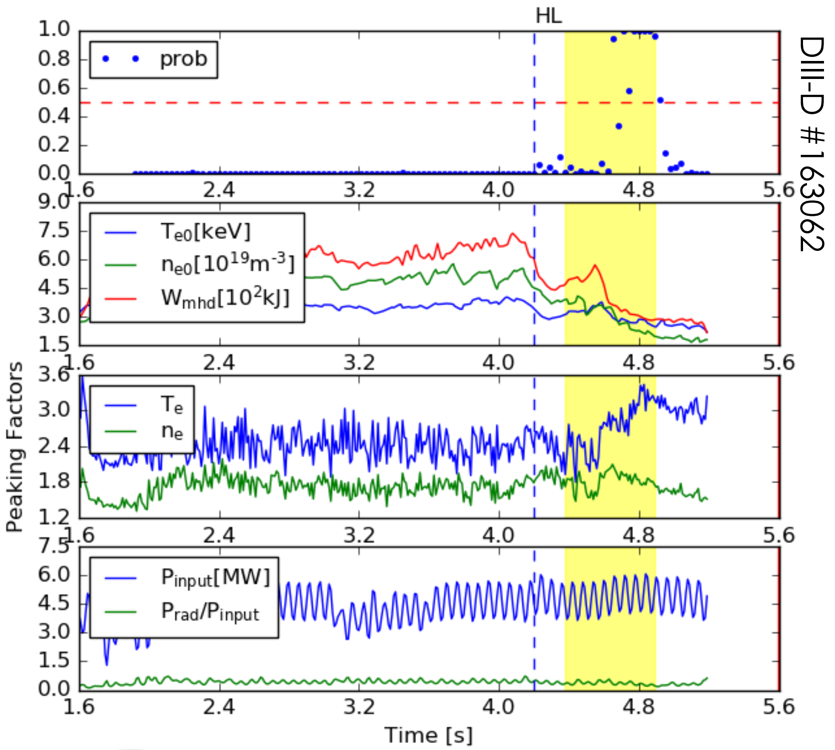
Metal Ring Campaign | ELM pacing | Initial Label



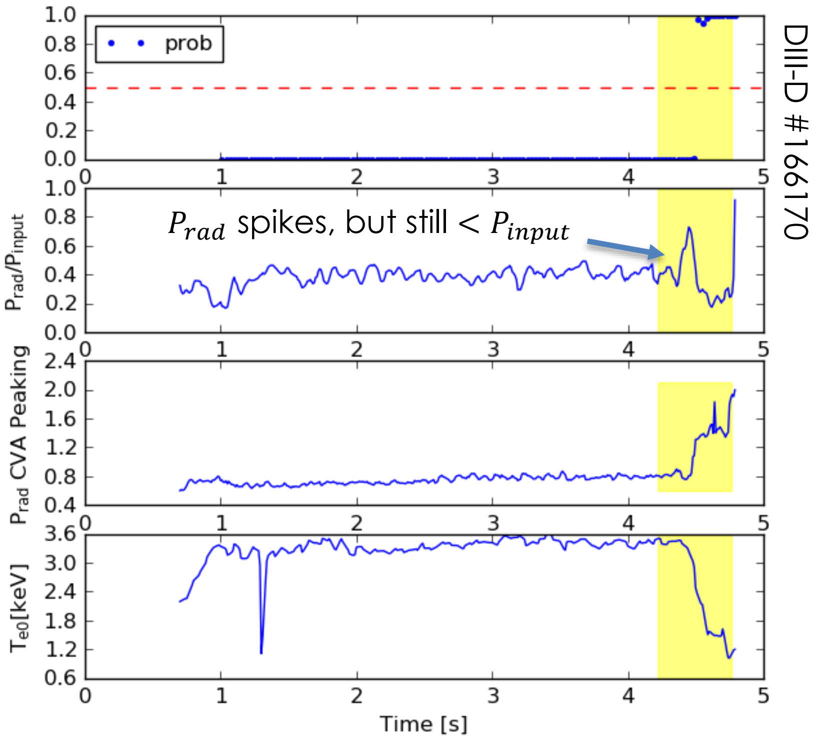
What went wrong? Common causes of missed detections & false positives

- **Early/late detections can cause missed warnings (classified as false negatives here)**
- **Several false positives due to strict event definitions (inflates FPR estimate)**
  - Better features (a binary threshold signal for $P_{rad} > P_{input}$, in the case below) can help

Late Detection (H-L transition)



False Positive (core radiative collapse)



Some other things to worry about …

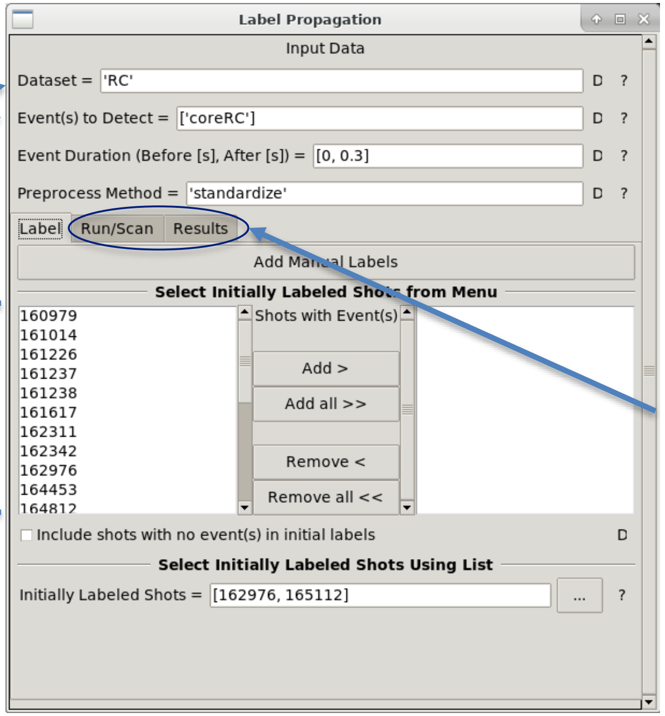
- **Curse of Dimensionality**
  - Recall each sequence $\vec{x}_i$ lives in $(N \times T)$-dimensional space, where N is the number of input signals and T is the number of times steps per sequence
  - Applications in this presentation were in 42, 24, and 18 dimensional spaces
  - Adding time resolution or signals could eventually make the problem intractable
- **Aliasing**
  - Number of time steps T should cover typical frequency of each input signal
  - For this application, several signals $(P_{input}, B_r^{n=1})$ were filtered
- **Kernels**
  - For this application, the radial basis function was used to find neighboring sequences
  - Other kernel functions may be better suited (think k-nearest neighbors, KNN)

OMFIT Module in Development for Arbitrary Event Detection

- **Module workflow facilitates using label propagation to quickly build a database**
  - label events, execute label propagation runs & parameter scans, and validate results

User can import arbitrary dataset for arbitrary event

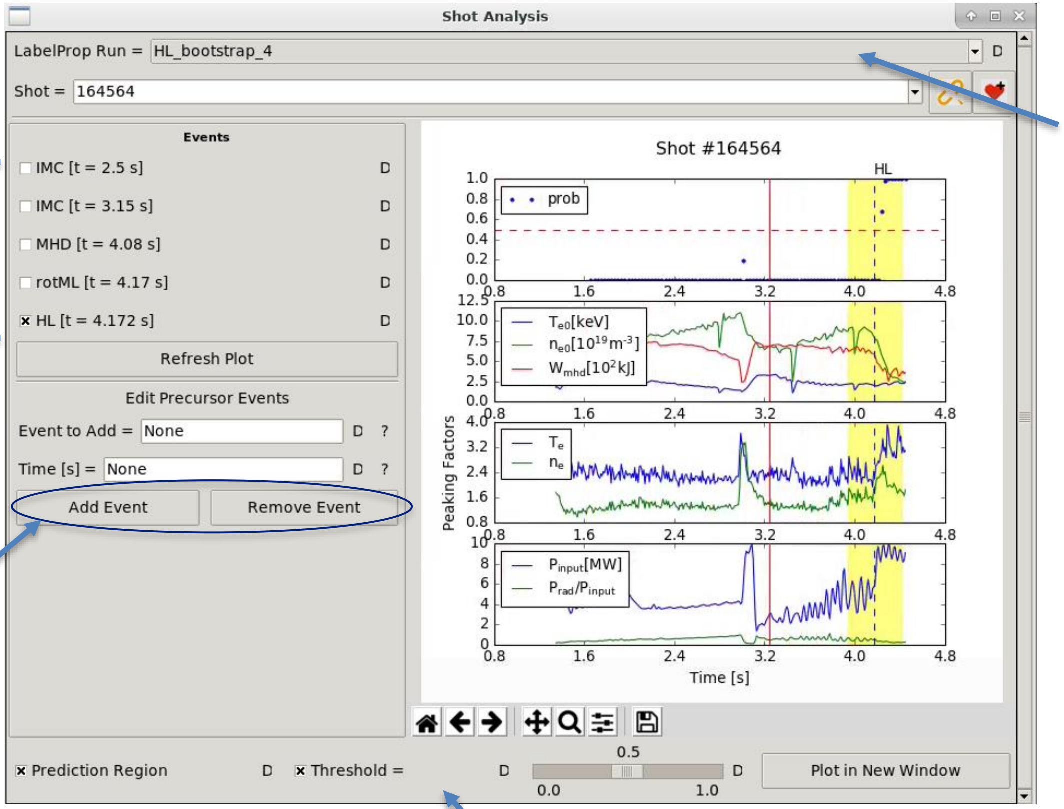
Choose shots to initially label



View chain of events for any shot

Generate customized label prop graph & view performance results

Add/remove verified events to/from the database



Compare predictions from many versions of algorithm

Change threshold for more confident predictions

DIII-D NATIONAL FUSION FACILITY

DISCLAIMER

DIII-D NATIONAL FUSION FACILITY