Machine Learning Studies of JET Disruptions

W. Tang, M. Parsons, E. Feibush, J. Kates-Harbeck, K. Felker
PPPL, Princeton University
J. Choi, Oak Ridge National Laboratory
T. Kurc, Stony Brook University
A. Murari, J. Vega, A. Pereira, I. Nunes, EuroFusion JET

IEA Workshop: Theory & Simulation of Disruptions
Princeton Plasma Physics Laboratory
Princeton, NJ

July 20-22, 2016

SITUATION ANALYSIS

Most critical problem for MFE: *avoid/mitigate large-scale major disruptions*

- Approach: Use of big-data-driven statistical/machine-learning predictions for the occurrence of disruptions in JET
- Current Status: ~ 6+ years of R&D results (led by JET) using SVM-based ML on *zero-D* time trace data executed on modern clusters yielding ~ *reported success rates ranging from 80 up to 90% for JET, BUT > 98% with false alarm rate < 3% actually needed for ITER (Reference – P. DeVries, et al., June 2015)*
- **PPPL Team Goals include:**

(i) improve physics fidelity via development of new *ML multi-D, time-dependent software including better classifiers;*

(ii) develop *"portable"* predictive software beyond JET to other devices and eventually ITER; and

(iii) enhance execution speed of disruption analysis for very large datasets via *development & deployment of advanced ML software via* ***SVM (Support Vector Machine) & DRNN (Deep Recurrent Neural Network)*** *methods*

NOTE: → *EUROfusion JET leadership has provided PPPL with collaborative access to its large disruption-relevant data base, including multi-D features*

CLASSIFICATION

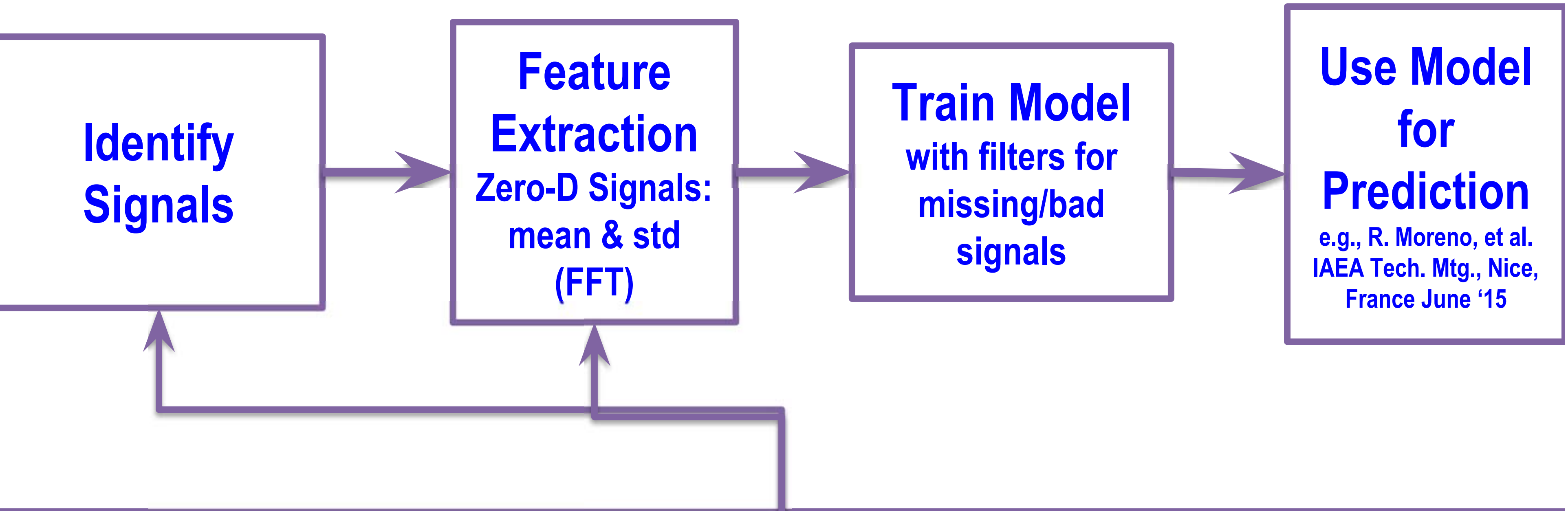
- Binary Classification Problem:
  - *Shots are Disruptive or Non-Disruptive*
- Supervised ML techniques:
  - Physics domain scientists combine knowledge base of observationally validated information with advanced statistical/ML predictive methods. Shots can be labeled D/ND retrospectively.
- Machine Learning (ML) Methods Engaged:

  Basic SVM approach initiated by JET team leading to APODIS software and *later to DPFD software @ PPPL*; and

  New Deep Learning Recurrent Neural Net (DRNN) approach @ PPPL
- Approach: (i) *examine normalized/dimensionless data;* (ii) *use training set to generate model;* (iii) *use trained model to classify new samples*

  ⟶ Targeted multi-D data analysis will require new signal representations

ML SVM Workflow



Moving forward, PPPL team will focus on *multi-dimensional (instead of present zero-D time trace) signals*

e.g. radial temperature profile

- ***Many interesting possibilities for more informative, physics-motivated features***
- *For SVM: Effective feature extraction is key challenge, can't just feed raw data*

SVM Approach

- **14 Feature vectors** are extracted from raw time series data

  7 signals* x 2 representations+

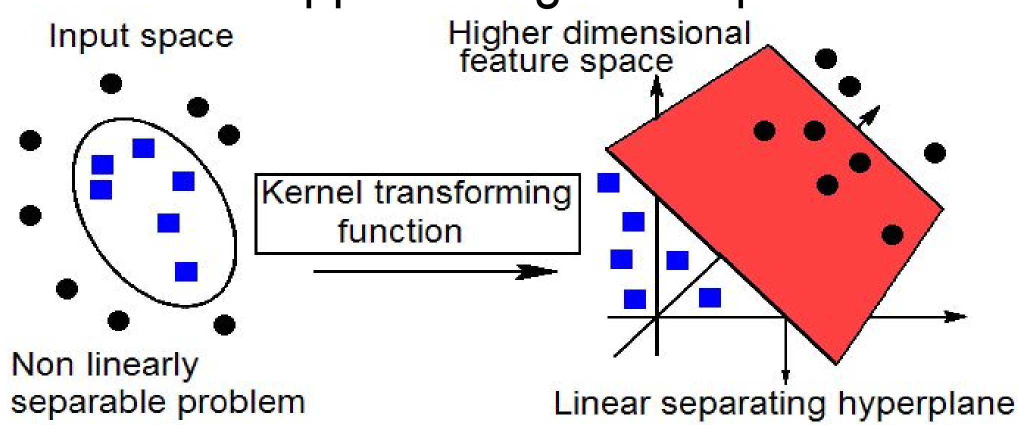
***Signals:**

1. Plasma current [A]
2. Mode lock amplitude [T]
3. Plasma density [m^{-3}]
4. Radiated power [W]
5. Total input power [W]
6. d/dt Stored Diamagnetic Energy [W]
7. Plasma Internal Inductance

+Representations (set of 32 samples at 1 kHz):

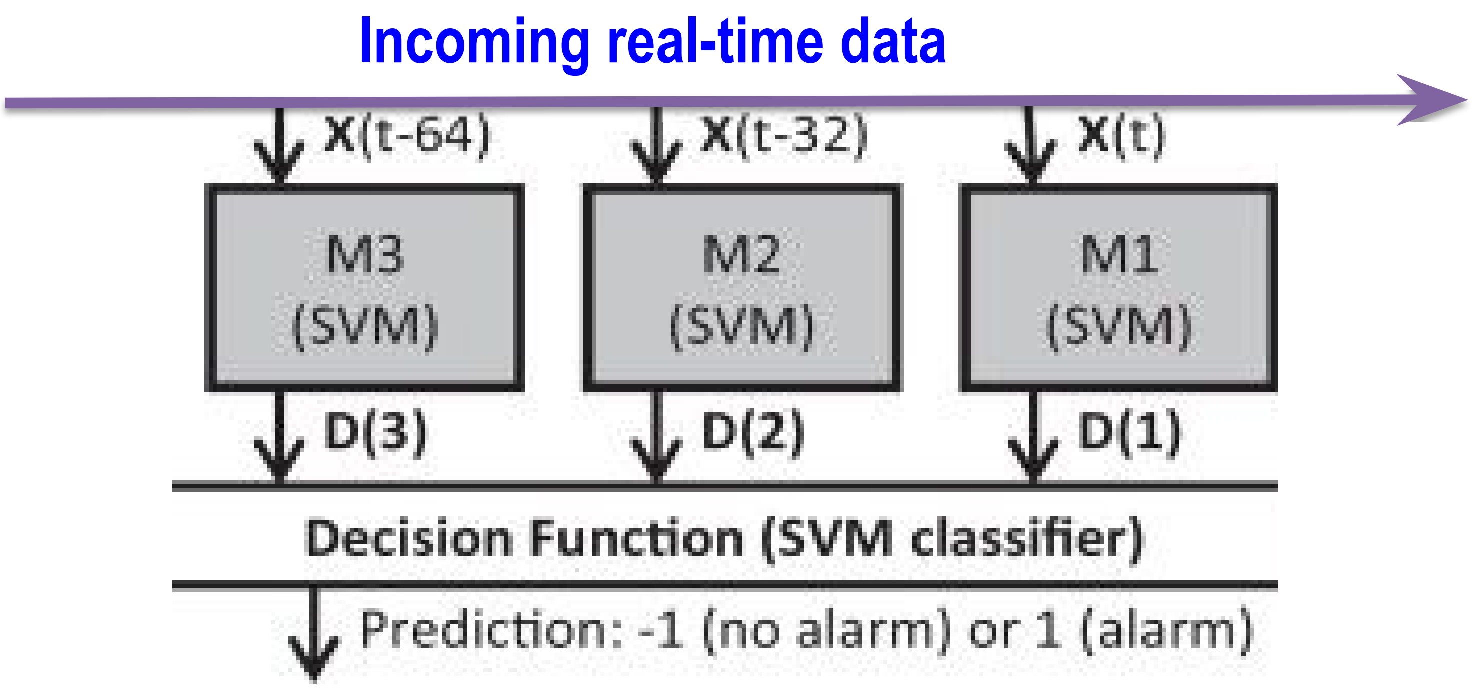
1. Mean
2. Standard deviation of positive FFT spectrum (excluding first component)

- Feature vectors are remapped to higher-D space



APODIS: Multi-tiered SVM

→ ***separate SVM models trained for separate consecutive time intervals preceding disruption***
- *applied to APODIS Code developed by J. Vega, et al.*



Reference: J. Vega *et al. Fusion Engineering and Design*, **88** (2013)
[APODIS → "Advanced Predictor of Disruptions"]

SVM R&D at PPPL

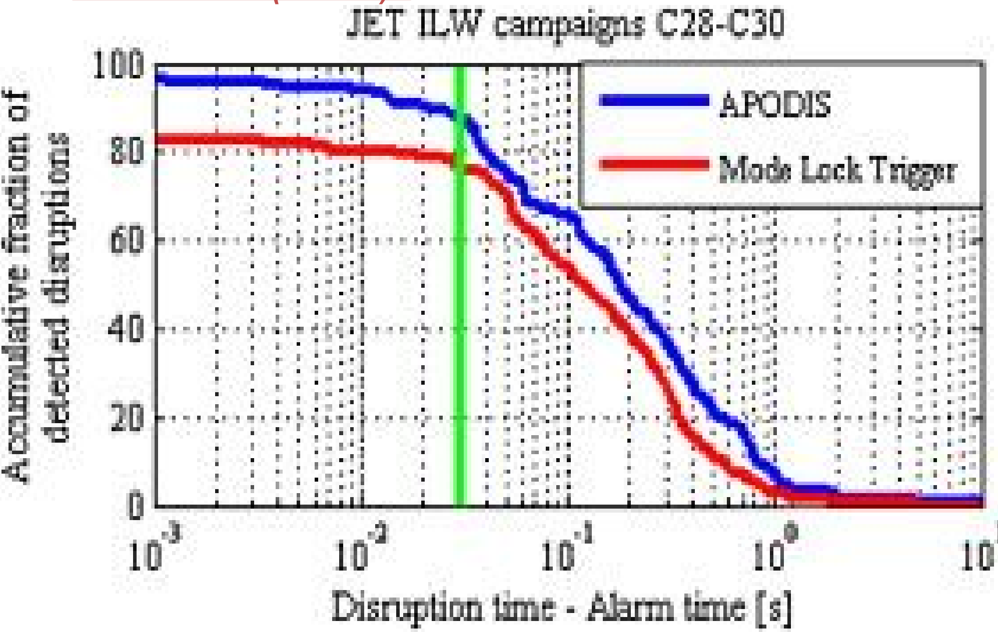
- E. Feibush, PPPL → utilized access to JET database to master navigation of MDS+ tree; extracted relevant signals, video, etc.
  Acquired and implemented SVM framework from APODIS for benchmarking PPPL software development → producing "DPFD"

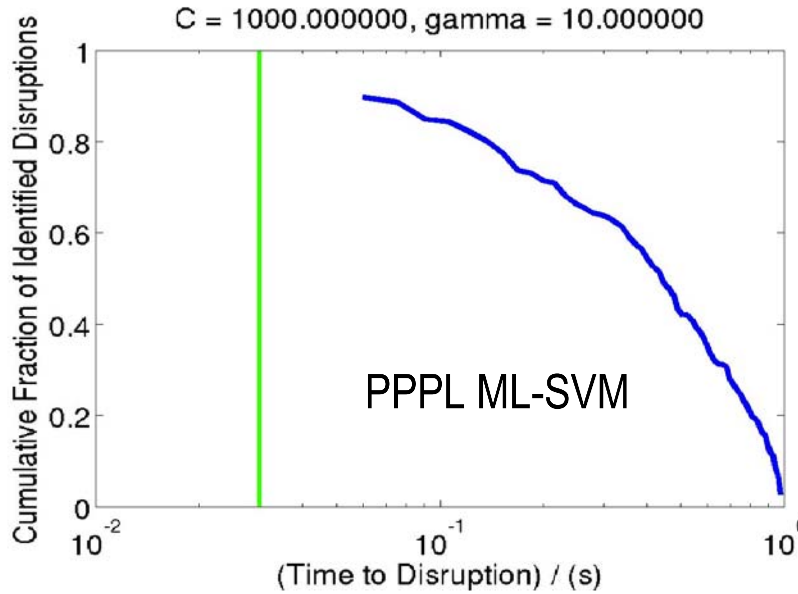
- M. Parsons, PPPL → Rewrote development framework to be self-contained within Matlab 2014, resulting in "DPFD" software:
  -- Simple training / testing possible in ~1 min
  -- DPFD "Starter Kit" with documentation now available for interested users

RESULTS for JET ILW Disruption Data: Benchmarking PPPL SVM DPFD Analysis with APODIS

30 ms before disruption *(Ref.-- P. DeVries – ITER disruption prediction requirements → mitigation trigger time > 30 ms)*

Both APODIS and DPFD give ~ 90.0% detected disruptions, FP rate lower for APODIS (~1%).





- APODIS trained on 738 disruptive and 2,035,000 non-disruptive samples

- PPPL's DPFD trained on 975 disruptive and 975 non-disruptive samples

ML Predictive Challenges & Opportunities

- CHALLENGE: SVM results at JET currently delivers in range of 80% to 90% accuracy ***30 ms before disruption***, but *need to consistently achieve Actual Goal for reliable ITER operation → 98% with acceptable false positive rate*
- PPPL FOCUS: Improve predictive performance by enhancing physics fidelity following "*Supervised ML via SVM and Deep Learning RNN*

→ *Function of time + spatial dimensions, including profile/gradient information*

-- *Serve as inputs for regression-type formulation of stability thresholds for improving physics-based classifiers*

-- *Investigate parametric scaling trends (complementary to e.g., S. Sabbagh's approach)*

-- *Explore inclusion of threshold conditions for key disruption precursors (currently NOT included in SVM classifiers) – e.g., Neoclassical Tearing Modes (NTMs)*

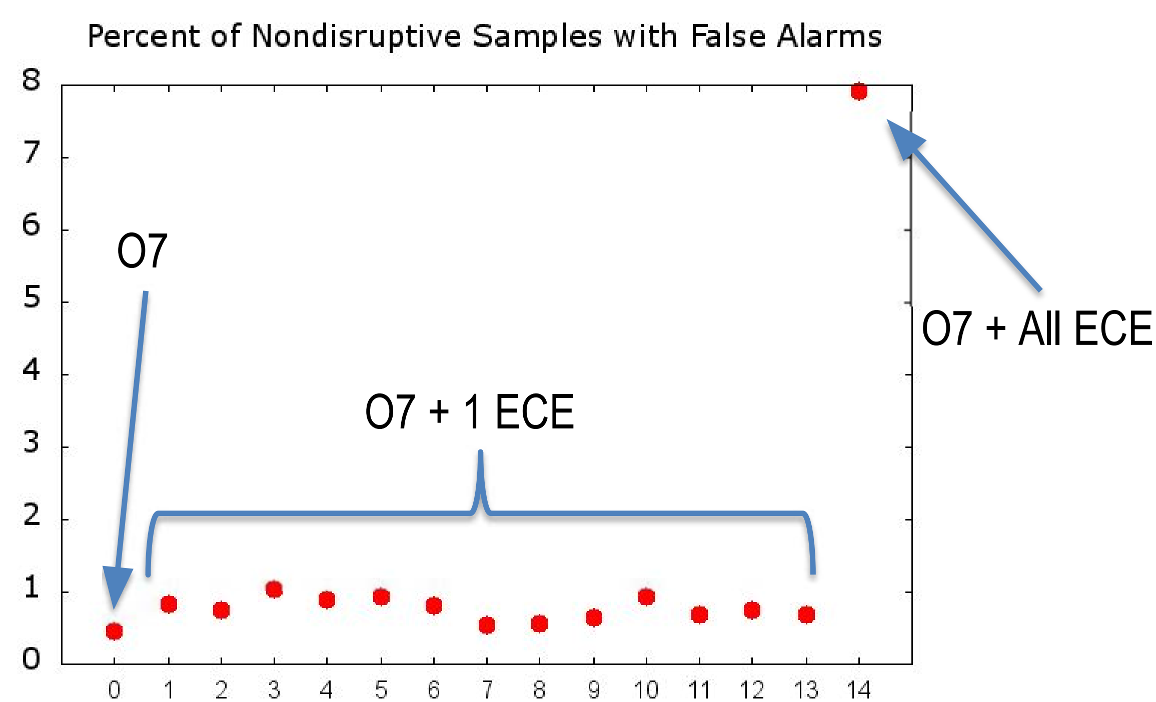
Initial Analysis of 1-D ECE Profiles (M. Parsons)

Work directly with T_e profiles (not raw data):

- ~98% of shots had data spanning R = [2.9, 3.5] m (B_0 = 2.96 m, a ~ 1.25 m,no edge coverage)
  - Take T_e at 13 radial positions (every 5 cm in domain)
- Adding 13 ECE channels to O7 signals causes sparsity, loss of statistical significance & adds false alarms. No meaningful improvement to prediction success

Early "lesson learned":

- SVM: Be very selective in adding features to the Original 7 ("O7") to avoid "Curse of Dimensionality"



Initial Cross-Machine Analysis

Approach: Train model with JET (CFC) data, Test on NSTX data

- Using JET normalizations: 10 / 37 detected (all within 64 - 96 ms of warning)
- Move to portable, **physics-based normalization:**

1. Normalized Plasma Current (I_N)	$I_p/(aB_T)$
2. Mode Lock Amplitude Fraction (f_{ML})	MLA/B_T
3. Plasma Internal Inductance (l_i)	l_i
4. Greenwald Density Fraction (f_{gw})	$n/(I_p/\pi a^2)$
5. Radiated Power Fraction (f_{rad})	$P_{rad}/(P_{in} - \dot{W}_{dia})$

- Shows ***100% success in disruption prediction*** over wide range of SVM parameters (258/258 shots)
- All NSTX shots disrupt: inability to assess false alarms motivates testing DIII-D data as next immediate target and also NSTX-U
- Productive recent discussions → strong interest at GA (R. Nazikian, T. Strait, B. Tobias, N. Logan, …. + E. Koleman at Princeton U/PPPL)

Recent Developments & Further Challenges & Opportunities

- Two 2016 DOE CSGF Fellows on ML Project @ PPPL: Julian Kates-Harbeck (Harvard U.) and Kyle Felker (Princeton U.)

- Matthew Parsons to move to ITER on Fullbright Fellowship (Fall, 2016)

- Portability Challenge for ML Predictive Software (beyond JET):
  - First investigate applicability to NSTX/NSTX-U disruption database, and then move on to others (DIII-D, ASDEX, EAST, KSTAR, …..) → *ultimately leading to ITER*

- Alternative Methods to SVM: Generative Topographic Mapping (GTM) with Deterministic Annealing (collaboratively with J. Choi of ORNL)
  - Can be used in parallel with SVM or used to improve choice of classifiers as part of SVM workflow

Projecting 14D space (7 signals x 2) into 3D space to develop SVM Classifers → movie shows separability of Disruptive from Non-Disruptive Data:

DA (Deterministic Annealing) Method

Jong Choi, ORNL

Generative Topographic Mapping (GTP) using Deterministic Annealing (DA)

References:

- J. Y. Choi, et al. Science Direct, Proc. Computer Science 00, 1-10 (2010)

- Geoffrey Fox, et al., Parallel Processing Letters, May 17, 2013.

DEEP LEARNING RECURSIVE NEURAL NETS (RNN) APPROACH

Julian Kates-Harbeck, CSGF Fellow from Harvard U.

→ ***Rapid development of new GPU-compatible predictive software with results benchmarked vs. those from SVM analysis***

Very Promising Approach to Analysis of Higher Dimensional Signals via Deep Learning RNN

1D:

(i) radial temperature profiles; (ii) density profiles; & (iii) radiation profiles

Goals:

-- Capture ***more physics and improve predictive capability***

-- Efficiently address challenges of more data and longer training time

→ *modern HPC training (e.g., via GPUs & MPI) progressing rapidly!*

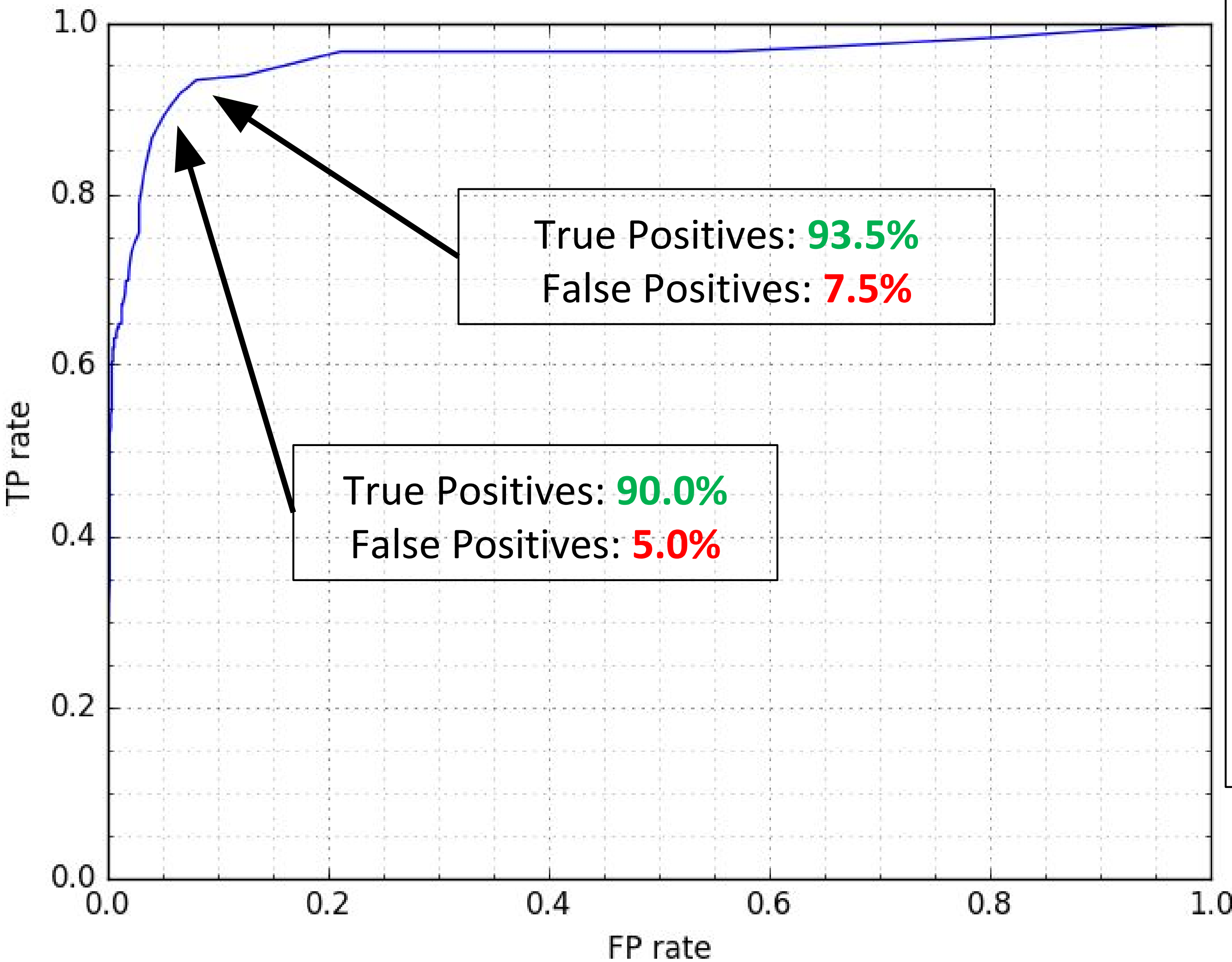
-- Demonstrate how Neural Networks can extract salient features from higher-D data dimensional data automatically

-- Demonstrate improvements in accuracy of ML predictions

→ including harvesting new physical insights in timely way

RNNs Performance

Performance Tradeoff: Tune True Positives (good: correctly caught disruption) **vs. False Positives** (bad: safe shot incorrectly labeled disruptive).



RNN Data:

- Testing **1200 shots** from Jet ILW campaigns (C28-C30)
- **All shots used**, no signal filtering or removal of shots

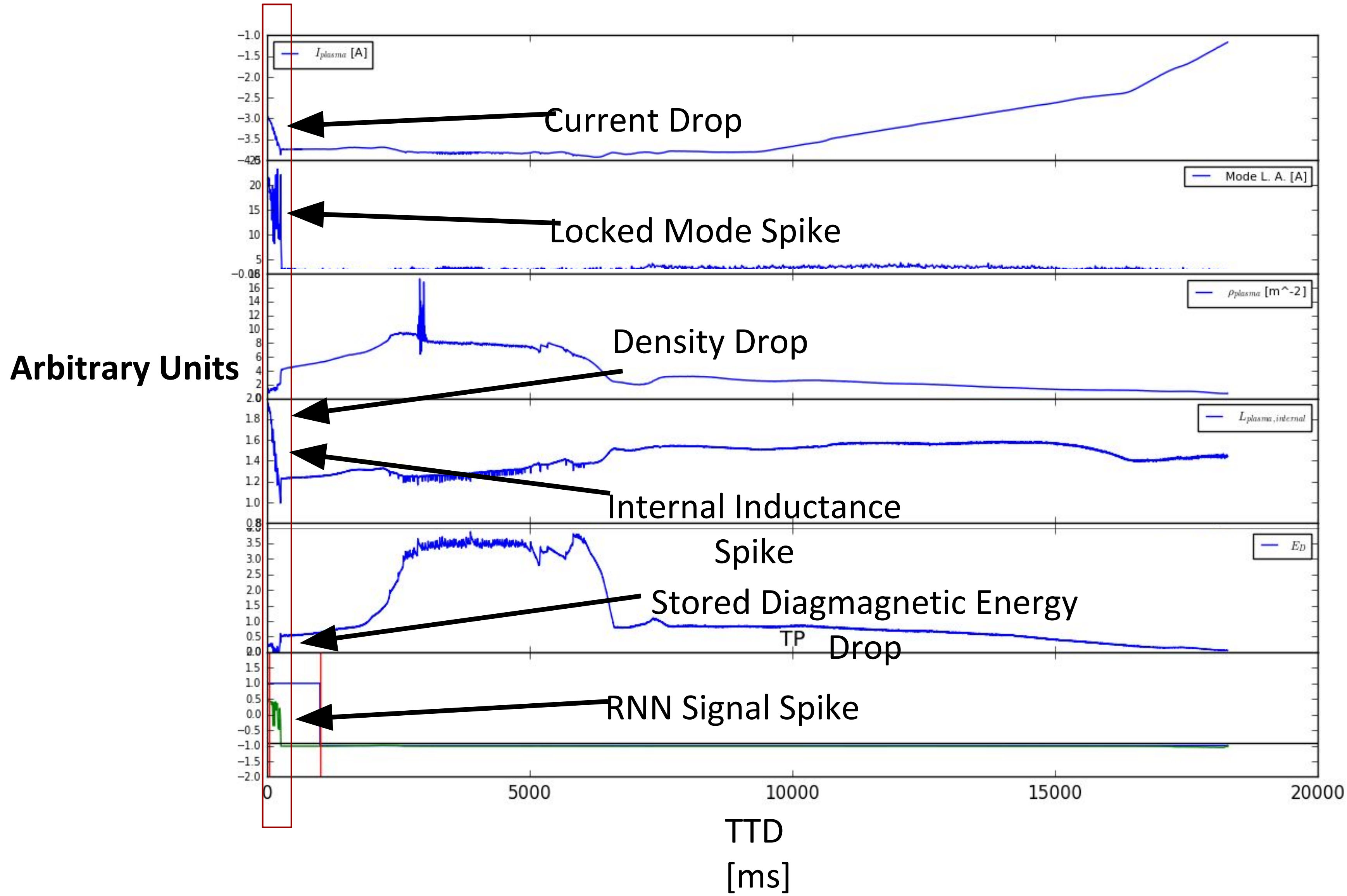
Jet SVM work:

- **990 shots** from same campaigns
- **Filtering** of signals, **ad hoc removal of shots** with abnormal signals
- **TP 90%, FP 1%**

Vega, Jesús, et al. "Results of the JET real-time disruption predictor in the ITER-like wall campaigns." *Fusion Engineering and Design* 88.6 (2013): 1228-1231.

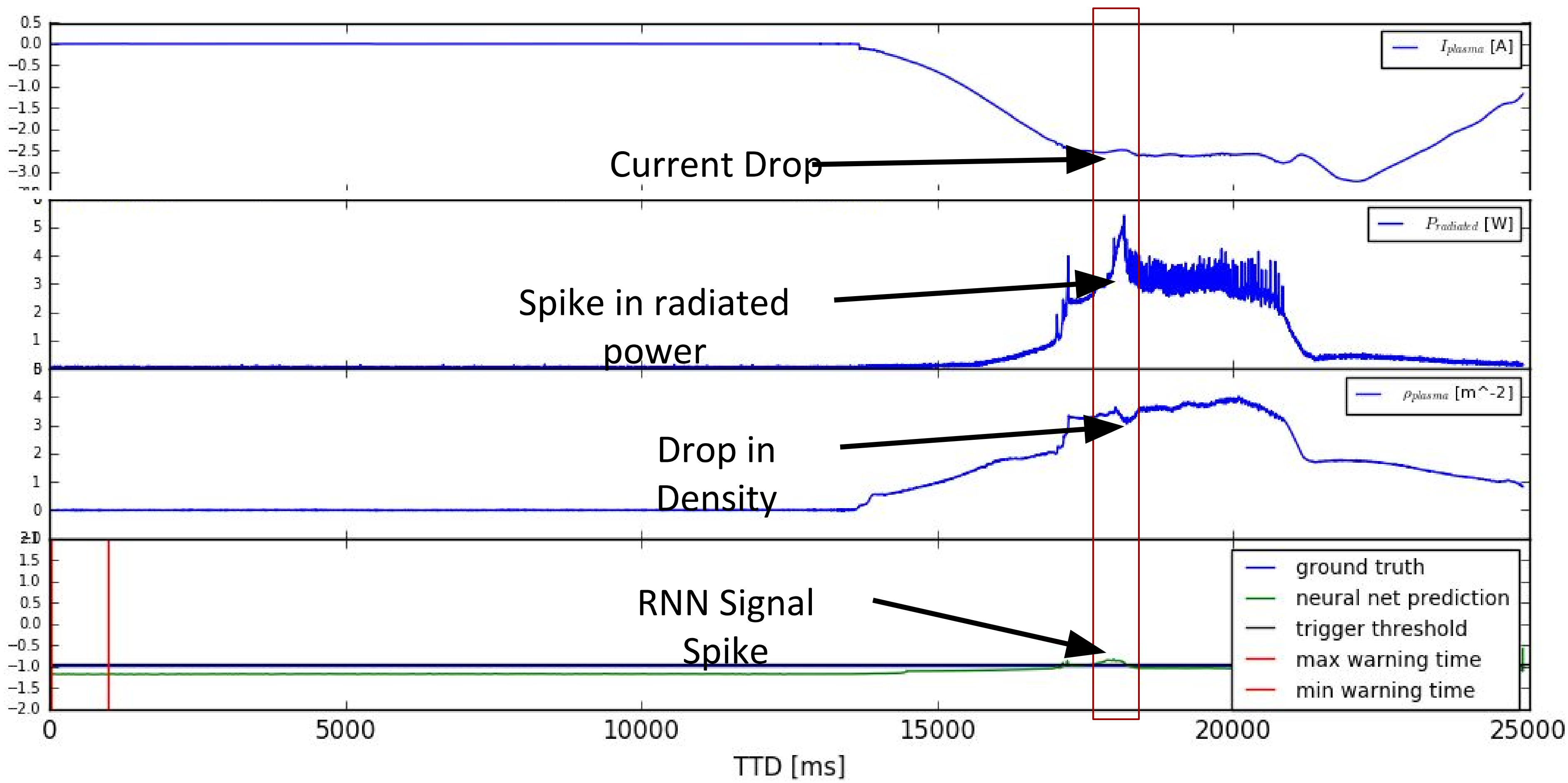
RNNs Predictions

True positive example



RNNs Predictions

False Positive Example or “Minor” Disruption?



RNN Data Challenges & Innovative Opportunities

- **Hyperparameter Tuning Challenges**
  including addressing size of hidden layers; number of layers; learning rate, initialization, etc.
  Approach: Try many models and choose the best
  → *Need rapid training time* to iterate!

- **Larger Data Set Challenges**
  For 0D: > 40GB and for 1D >1TB)
  → *Need rapid training time* to process all data!

- **Signal Normalization & Outlier Detection Challenges**
  Try many approaches and choose the best.
  → *Need rapid training time* to iterate

HPC engagement (e.g. via GPUs and MPI) is key and looks very promising

RNNs: HPC Innovations Engaged

GPU training

Neural networks use dense tensor manipulations, efficient use of GPU FLOPS

Over 10x speedup over multicore node training

Distributed Training via MPI

Linear scaling:

- Key benchmark of “time to accuracy”: we can train a **model that achieves the same results nearly N times faster with N GPUs**

Scalable

- **to 100s or 1000s of GPUs/ Leadership Class Facilities!**
- **>> TBs of data or more**
- Example: Best model training time on full dataset (~40GB, 4500 shots) of 0D signals training
  - **SVM** (JET) **: > 24hrs**
  - **RNN** (distributed, 20 GPU’s) : **~40min**

Deep Recurrent Neural Networks (RNNs): Basic Description

- **“Deep”**
  - Hierarchical representation of complex data, building up salient features automatically
  - Obviating the need for hand tuning, feature engineering, and feature selection
- **“Recurrent”**
  - Natural notion of time and memory
  - At every timestep, output depends on
    - Last Internal state “s(t-1)” ***Recurrence!***
    - Current input x(t)
  - The internal state can act as memory and accumulate information of what has happened in the past

Internal State (“memory/ context”)

$$s_t = A(W_{in}x_t + W_{recurrent}s_{t-1} + b)$$
$$h_t = W_{out}s_t$$

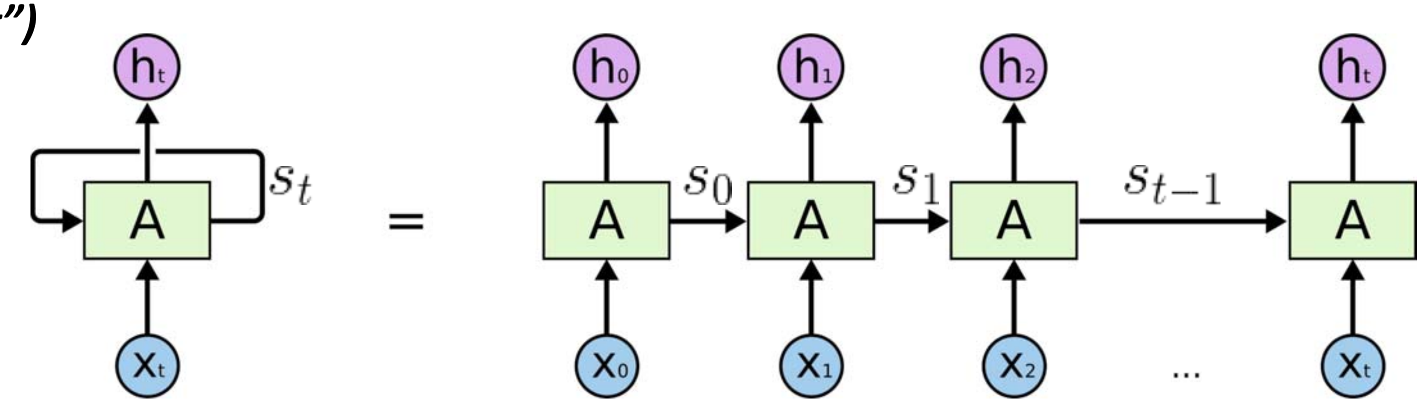
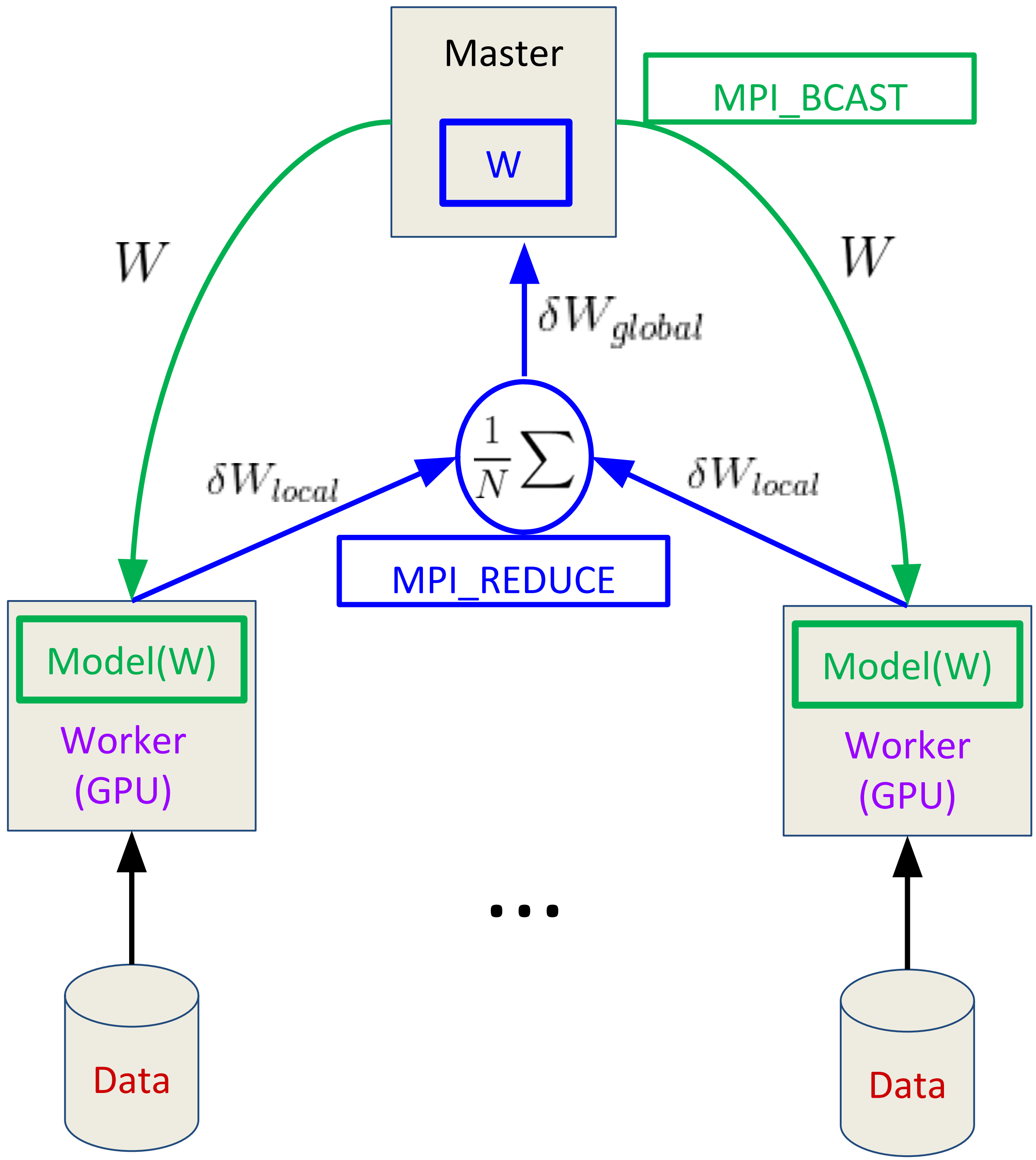


Image adapted from: colah.github.io

Distributed Training

Distributed Training

- Keep N model instances
- Each computes a gradient step on a different subset of the data
- The gradients are ***reduced*** (averaged) and the updates are made to a global set of parameters
- The global parameters are ***broadcast*** to the N models
- Efficient communication using custom **MPI** implementation



Summary

- **Fusion Energy Mission Relevance**:

  -- *Goal of Magnetic Fusion Energy goal is demonstrating the scientific & technical feasiblity of delivering Fusion Power*

  -- *Most critical associated problem is to avoid/mitigate large-scale major disruptions*

- **Relevance to HPC:**

  -- ***New focus*** *on development of large-data-driven* ***"machine-learning" statistical methods – Support Vector Machines (SVMs) and Deep Learning Recurrent Neural Nets (RNNs) --*** *as an exciting alternative/complement for conventional "hypothesis-driven/first principles" predictive methods*

- **Associated Challenge:**

  → *Significant improvements over zero-D SVM-based machine-learning capabilities to achieve ~ 98% success rate with portability of software to ITER via enhanced physics fidelity (capturing multi-D) with improvement in execution time enabled by access to advanced HPC hardware (e.g., large GPU systems).*

Additional Materials

Disruption Data Subset

# Shots	Disruptive	Nondisruptive	Totals
Carbon Wall	324	4029	4353
Beryllium Wall	185	1036	1221
Totals	509	5065	5574

Sample Signals (0-D time trace)	Data Size (GB)
Plasma Current	1.8
Mode Lock Amplitude	1.8
Plasma Density	7.8
Radiated Power	30.0
Total Input Power	3.0
d/dt Stored Diamagnetic Energy	2.9
Plasma Internal Inductance	3.0

JET produces ~ Terabyte (TB) of data per day

~55 GB data collected from each JET shot

→ ***Well over 350 TB total amount with multi-dimensional data yet to be analyzed***