Potential Quantum Computing Enhancement of Machine Learning in Predictive Fusion Energy Applications

William Tang
Princeton University, Princeton, NJ USA
&
Federico Spedalieri
Information Sciences Institute, University of Southern California

DOE ASCR Workshop on Quantum Computing for Science

Bethesda, Maryland

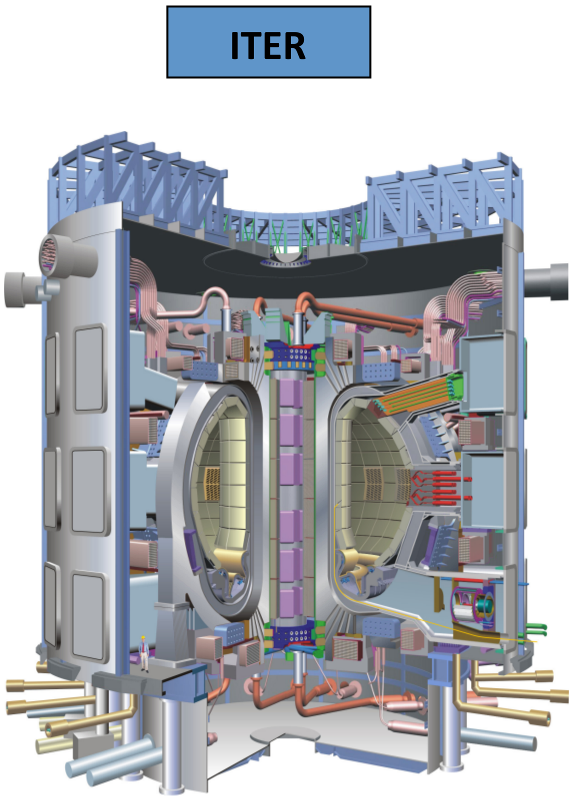
February 17-18, 2015

DOE FUSION ENERGY MISSION: *Demonstration of the Scientific and Technological Feasibility of Fusion Power*

- **ITER:** ***~$25B facility under construction in France***
  -- 7 governments representing over half of world
  → dramatic next-step for Magnetic Fusion Energy (MFE) producing a sustained burning plasma
  -- Today: 10 MW(th) for 1 second with gain ~1 [JET]
  -- ITER: 500 MW(th) for >400 seconds with gain >10

- **Ongoing R&D programs worldwide [experiments, theory, *HPC*, and technology]** *essential to provide growing knowledge base for ITER operation targeted for ~ 2025*

→ Reliable HPC-enabled predictive capabilities required to cost-effectively plan, "steer," & harvest key information from expensive (~$1M/long-pulse) shots



US/EU Statistical Disruption Studies on JET [Joint European Torus]

Situation Analysis:

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

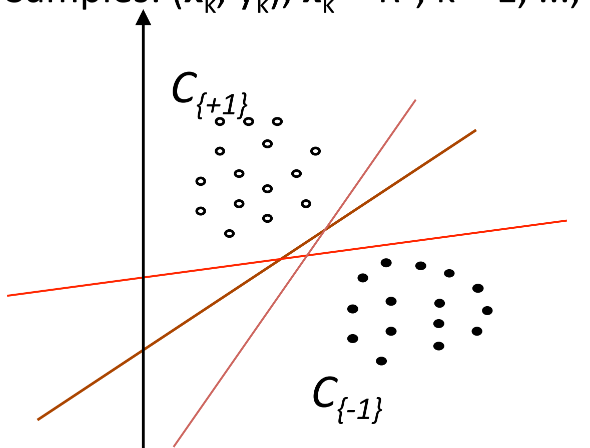
-- Conventional *"hypothesis-driven" MHD codes currently far away from achieving predictive capability needed for disruption avoidance in JET* → only experiment that achieved near "break-even" fusion energy production.

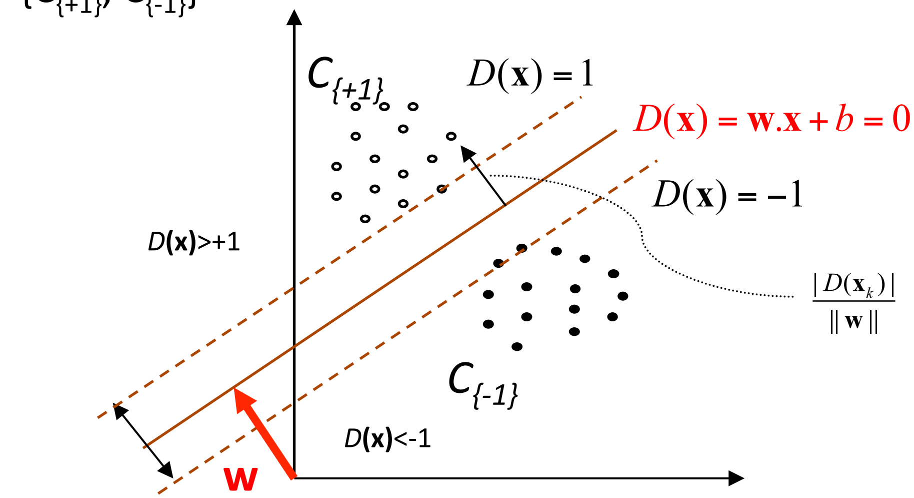
- Approach: Use of of large- data-driven statistical/machine-learning predictions for the occurrence of disruptions in JET
- Current Status: ~ 6 years of R&D results using SVM-based ML on *zero-D* time trace data executed on modern clusters yielding ***~ 80% success rate, BUT > 95% actually needed !***
- **Goal:** improve (i) physics fidelity via new *ML multi-D, time-dependent software* and (ii) execution speed via deployment of improved ML software on LCF's or ***possibly on innovative Quantum Computing systems*** appropriate for needed large-scale "data-mining" analysis of JET data

NOTE: → *JET has recently agreed to provide unique access to its huge disruption-relevant multi-dimensional data base that has yet to be analyzed.*

Supervised Classifiers: SVM

- Binary classifier
- Finds the optimal separating hyper-plane between classes
- Samples: $(\mathbf{x}_k, y_k)$, $\mathbf{x}_k \in R^n$, $k = 1, \ldots, N$, $y \in \{C_{\{+1\}}, C_{\{-1\}}\}$





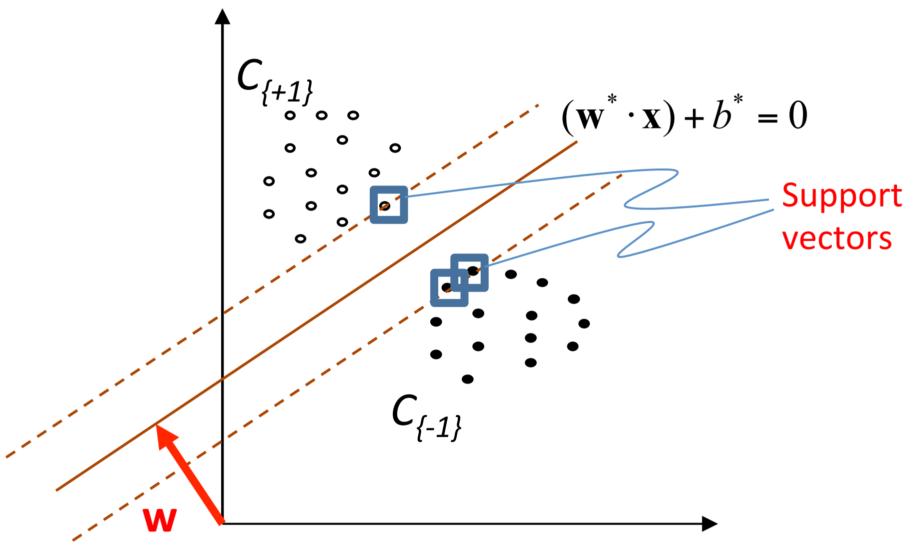
$$\frac{y_k D(\mathbf{x}_k)}{\|\mathbf{w}\|} \ge \tau, \quad y_k \in \{+1,-1\}, \quad k = 1,\ldots,N$$

- Find optimal hyper-plane by determining vector **w** that maximizes the margin τ
- To avoid infinite solutions due to presence of a scale factor: $\tau \|\mathbf{w}\| = 1$
- To maximize margin is equivalent to minimizing $\|\mathbf{w}\|$

i.e., Optimization problem: $\min_{\mathbf{w},w_0} \; J(\mathbf{w}) = \|\mathbf{w}\|^2, \quad \text{subject to} \quad y_k \left[(\mathbf{w}.\mathbf{x}_i) + w_0 \right] \ge 1$

Supervised Classifiers: SVM

$(\mathbf{x}_k, y_k)$, $\mathbf{x}_k \in R^n$, $k = 1, \ldots, N$, $y \in C_{\{+1\}}, C_{\{-1\}}\}$



- Solution: $$\mathbf{w}^* = \sum_{i=1}^{N} \alpha_i^* y_i \mathbf{x}_i$$

  α_i are the Lagrange multipliers

- Samples associated to $\alpha_i \neq 0$ are called "support vectors"

  $$\boxed{\mathbf{w} = \sum_{\text{support vectors}} \alpha_i^* y_i \mathbf{x}_i}$$

  The rest of training samples are irrelevant to classify new samples

- The constant b is obtained from any condition (Karush-Kuhn-Tucker)

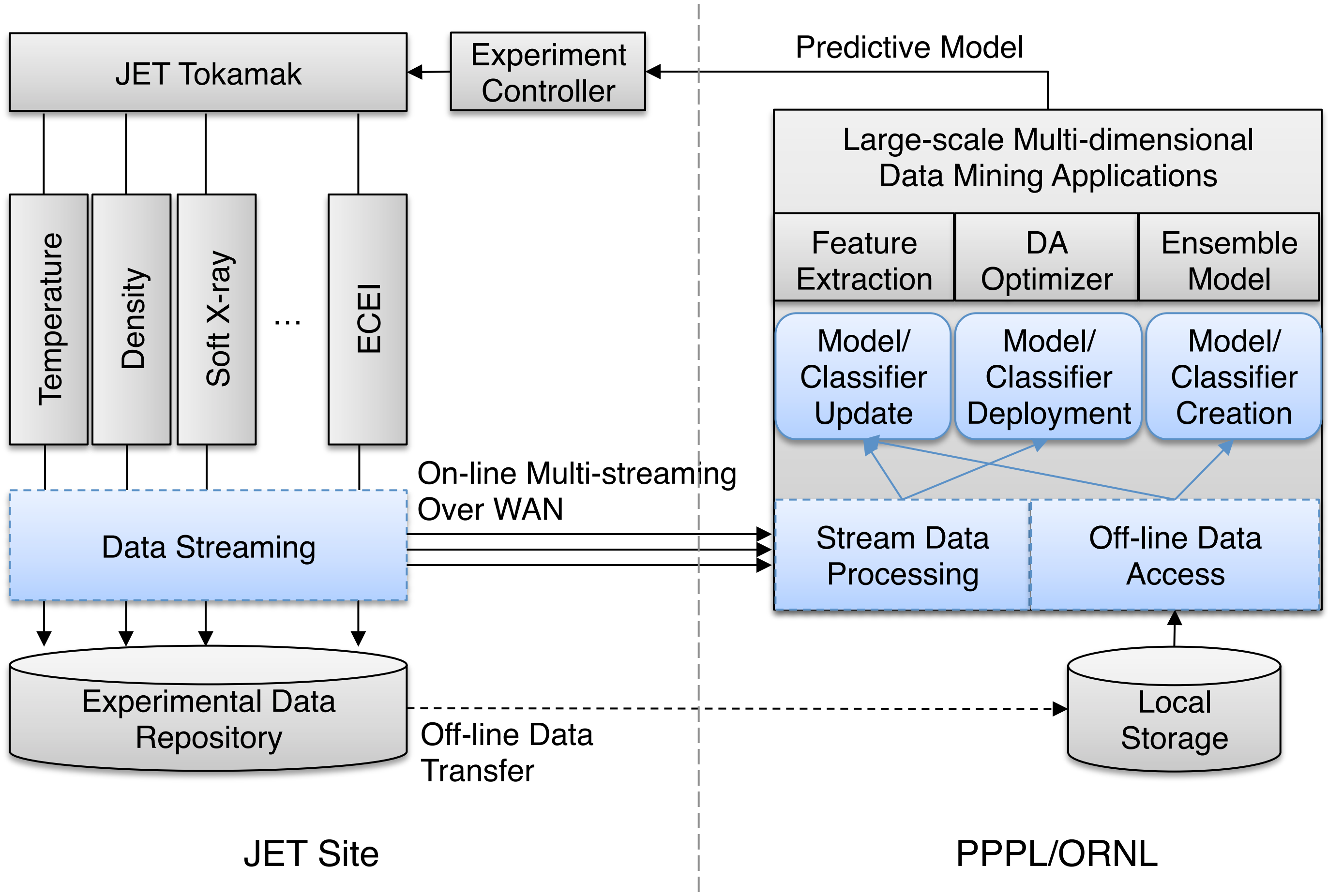
  $$\alpha_i \left[y_i \left((\mathbf{w} \cdot \mathbf{x}_i) + b \right) - 1 \right] = 0, \quad i = 1, \ldots, N$$

$D(\mathbf{x}) = \mathbf{w}^* \cdot \mathbf{x} + b^*$ is the distance (with sign) from **x** to the separating hyper-plane

Given **x** to classify

$$\textit{if sign}\left(\sum_{\textit{vectores soporte}} \alpha_i^* y_i (\mathbf{x} \cdot \mathbf{x}_i) + b^* \right) \geq 0, \quad \mathbf{x} \in C_{\{+1\}}. \quad \text{Otherwise } \mathbf{x} \in C_{\{-1\}}$$

Fusion Data Mining Diagram



NOTE: DA (Deterministic Annealing) Method

- "Generative Topographic Mapping by Deterministic Annealing," J. Y. Choi, et al. Science Direct, Proc. Computer Science 00, 1-10 (2010); Geoffrey Fox, et al., Parallel Processing Letters, May 17, 2013.

Machine Learning with Quantum Computers

Examples:

- Quantum Support Vector Machine [SVM]

<u>Ref.</u>: P. Rebentrost, M. Mohseni, & S. Lloyd Phys. Rev. Lett. 113, 130503 (2014)

Based on:
- *Fast quantum evaluation of inner products*
- *Fast quantum matrix inversion*

→ Designed for the quantum circuit model

- Training Strong Classifiers through Quantum Annealing

<u>Reference</u>: H. Neven, et al., arXiv:0912.0779 [quant-ph]; *http://en.wikipedia.org/wiki/Quantum_machine_learning*

$$\boxed{\vec{w}_{opt} = \arg\min_{\vec{w}} \left(\delta(\vec{w}) + \lambda' \|\vec{w}\|_0\right)}$$

<u>Main Task</u>: pick a set of weak classifiers out of a large library in order to minimize classification error

$$\delta(\vec{w}) = \left\| \vec{y} - \frac{1}{M}\vec{R}_{\vec{w}} \right\|^2 = \sum_{s=1}^{N_S} \left| y_s - \frac{1}{M}\sum_{i=1}^{M} w_i h_i(\vec{x}_s)\right|^2$$

Weight (0,1)

Label

Weak classifier

Training example

This is a quadratic unconstrained binary optimization (QUBO) that can be implemented on a quantum annealer

→ <u>*This has been run on the D-Wave device*</u>

Summary

• **DOE Mission Relevance**:

-- *Magnetic Fusion Energy with it's goal of demonstrating the scientific & technical feasiblity of delivering Fusion Power is an important DOE mission.*

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

• **Impact on Computing:**

-- *Development of large-data-driven* ***"machine-learning" statistical methods*** *as alternative/complement for conventional "hypothesis-driven/first principles" methods*

• **Challenges:**

-- *Needs significant improvements (from 80% to >95%) over zero-D SVM-based machine-learning capabilities with respect to physics fidelity (capturing multi-D) and execution time (moving beyond clusters to LCF's or* ***viable quantum computers****).*

→ Associated QC development challenge to produce ML software interface needed to connect to QC hardware (such as D-Wave)

• **Implications for Accelerating Scientific Knowledge Discovery:**

→ *Possible Quantum Computing impact via connection to Machine Learning Software*

• *Promising Approaches: (i) Quantum SVM designed for quantum circuit model; & (ii) Quantum Annealing for training strong classifiers*

→ *Exciting promise for stimulating progress in predicting complex behavior in DOE mission domains including example areas such as Fusion Energy.*