Transfer learning with a parsimonious disruption predictor: from JET C-wall to the metallic wall

J. Vega[1], J. D. Fernández[2], S. Dormido-Canto[2], R. Castro[1], A. Murari[3,4] and JET Contributors*

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK

[1]Laboratorio Nacional de Fusión, CIEMAT, Madrid, Spain
[2]Departamento de Informática y Automática, UNED, Madrid, Spain
[3]Consorzio RFX (CNR, ENEA, INFN, Universitá di Padova, Acciaierie Venete SpA), Padova, Italy
[4]Istituto per la Scienza e la Tecnologia dei Plasmi, CNR, Padova, Italy
* See the Appendix of F. Romanelli et al., Proceedings of the 25th IAEA Fusion Energy Conference 2014, Saint Petersburg, Russia

Abstract

The present contribution investigates the potential of the Learning Using Privileged Information (LUPI) paradigm for transfer learning. The goal is to develop a parsimonious disruption predictor with JET C-wall discharges and to perform transfer learning to ITER-like Wall (ILW) shots. Only a line integrated density (LID) signal is used for real-time predictions and the predictor is trained with the LID signal and the mode lock as privileged information. The database consists of 439 C-wall discharges together with 471 ILW discharges. An adaptive predictor is trained with the C-wall data and re-trainings are carried out after missed or tardy alarms. After processing the 439 C-wall discharges, the last predictor is applied to ILW shots. Again, the predictor is re-trained adding ILW data after missed or tardy alarms.

Objective

- The concept of transfer learning for disruption prediction means to train predictors with existing databases of at least one Tokamak and to apply the model to new devices (for example JET ➔ ITER) or after drastic changes in a Tokamak (for instance C-wall ➔ metallic wall)
- Nowadays, reliable transfer learning for disruption prediction is challenging[1]
- Previous attempts[2, 3, 4] have been carried out between AUG and JET
- The goal: development of parsimonious predictors with C-wall discharges and minimum number of signals

Supervised classifier

- Let's consider a training set whose only assumption about the examples is that they satisfy the *i.i.d.* hypothesis (independent and identically distributed samples)
- Let $\{z_1, z_2, \dots, z_n\}$ be the training set, where $z_i \in Z = X \times Y$ is a pair (x_i, y_i) consisting of the sample or feature vector x_i and its class y_i, where $y \in \{Y_1, Y_2, \dots Y_J\}$
- Given a new feature vector x, the objective of SVM[5] is to estimate its label y from the set of Y classes $y \in \{Y_1, Y_2, \dots Y_J\}$

SVM with privileged information[6] (SVM+)

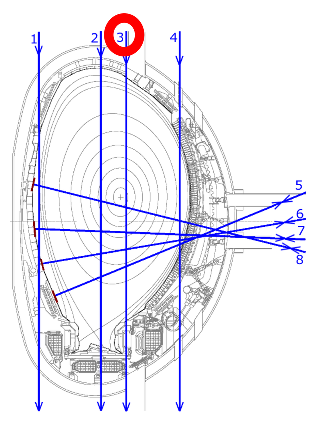
- There is information at training time that is not available at execution time
- Let's consider a training set of triplets

$$(x_i, x_i^*, y_i), x_i \in X, x_i^* \in X^*, y_i \in \{Y_1, Y_2, \dots, Y_J\}$$

- The only assumption is that they satisfy the *i.i.d.* hypothesis.
- X is the **decision space**: a space of vectors whose feature vectors contain information of past data
  - These features are available in real-time
- X* is the **correction space**: a space of vectors whose feature vectors contain information of past data
  - These features are not available in real-time
  - These data can be used at training time and it is known as privileged information
  - The privileged information helps in optimising the separation frontier (or decision function) in the decision space
- Given a new feature vector $x \in X$, the objective of SVM+ is to estimate its label y from the set of J classes $y \in \{Y_1, Y_2, \dots YJ\}$

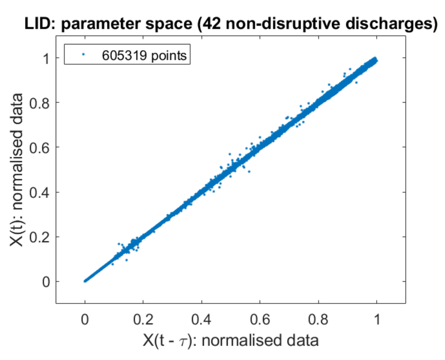
Database

- 439 C-wall discharges in the range 65988 – 68244
  - 30 unintentional disruptive shots
  - 409 non-disruptive shots
- 471 ILW discharges in the range 94152 - 95887
  - 79 unintentional disruptive shots
  - 392 non-disruptive shots
- Only discharges with plasma current above 2 MA are considered
- Only disruptions whose plasma current at disruption time is greater than 1.5 MA are taken into account



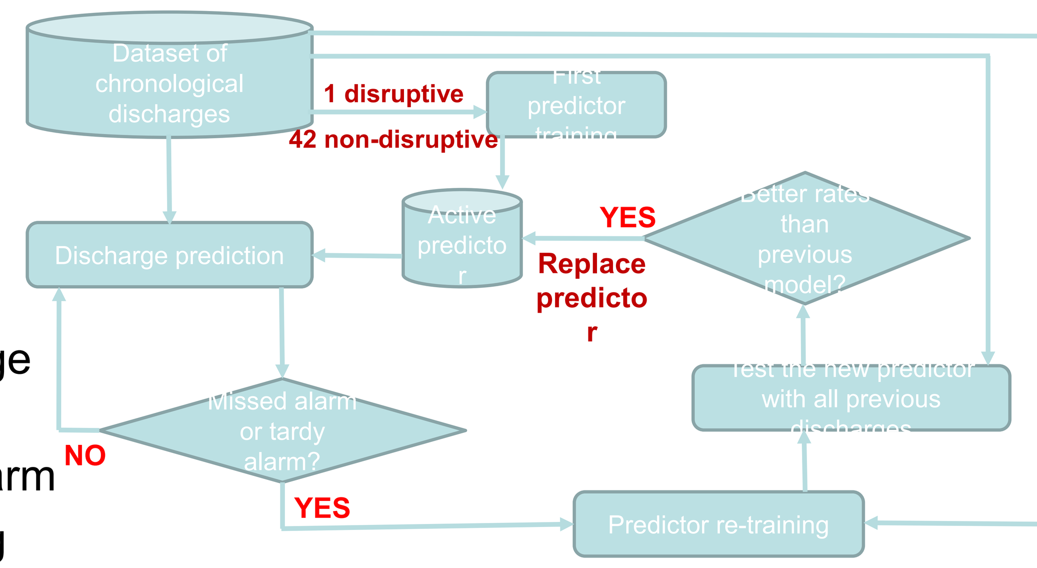
Parameter space of the predictor system

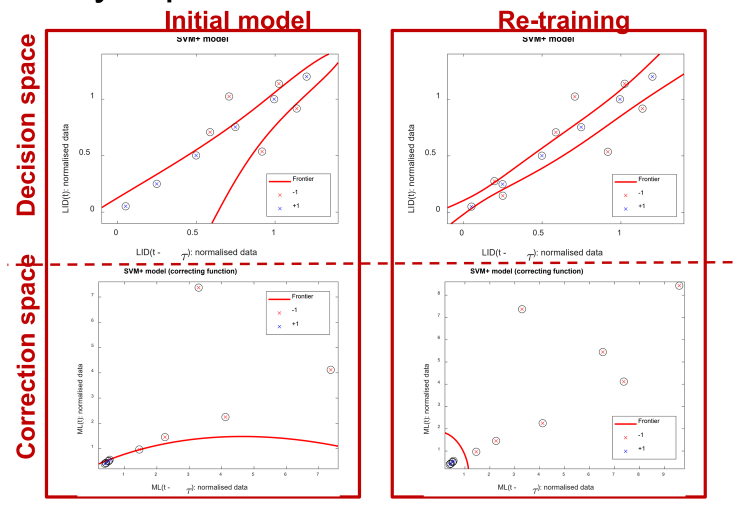
- A two-dimensional parameter space made up of the amplitudes of consecutive samples is used
  - This means to consider the deltas between consecutive samples
- This parameter space has been used to develop a real-time disruption predictor in JET based on centroids[7, 8]



C-wall discharges: adaptive predictor from scratch

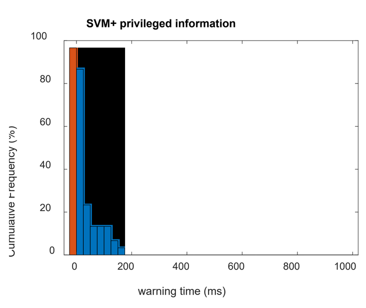
- Discharges are processed in chronological way
  - Only **LID signal** and **mode lock signal** (privileged information)
- First predictor with 1 disruptive discharge and 42 non-disruptive ones
- Re-training after a missed or a tardy alarm
- Only 2 predictors: 1st one + 1 re-training





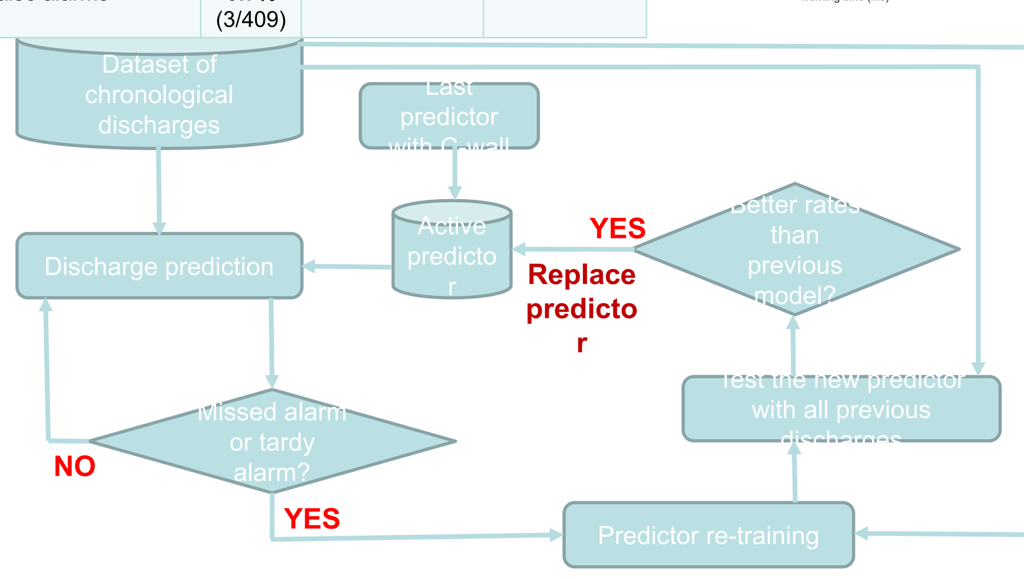
- 1 missed alarm, 3 tardy alarms and extremely reduced false alarms
- Only 1 re-training obtained better results with previous discharges
- All trainings provide separable models
  - Success rate with training samples is 100%

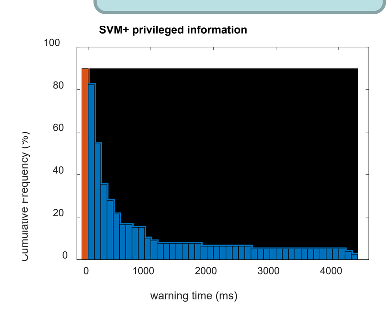
		Average warning time (ms)	Standard deviation (ms)
Success rate	**96.7%** (29/30)	-	-
Success rate with positive warning time	**86.7%** (26/30)	30	40
Tardy alarm	**10,0%** (3/30)	5	2
False alarms	**0.7%** (3/409)	-	-

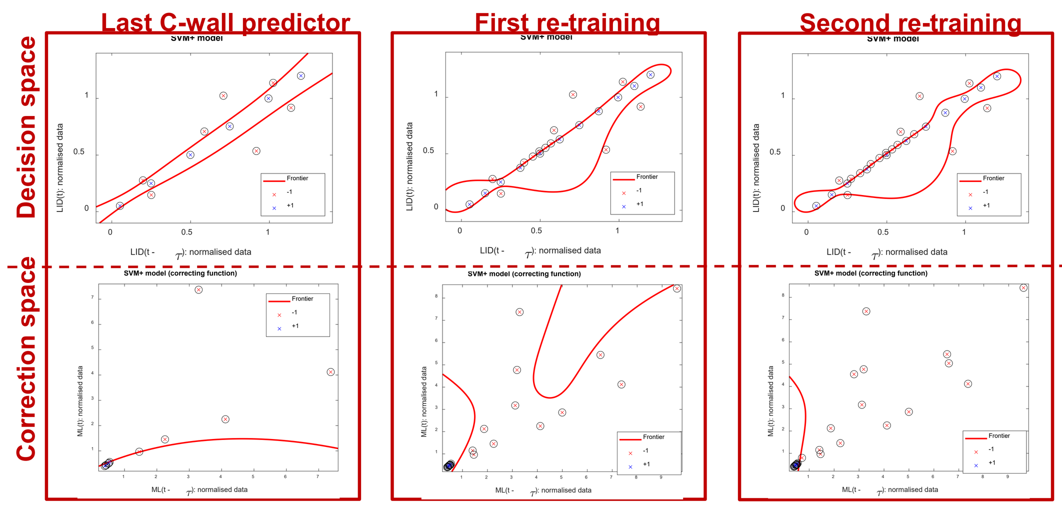


ILW discharges: adaptive predictor from transfer learning

- Discharges are processed in chronological way
  - Only **LID signal** and **mode lock signal** (privileged information)
- First predictor: the last one with C-wall data
- Re-training after a missed alarm or a tardy alarm
- Only 3 predictors: the last C-wall predictor + 2 re-trainings





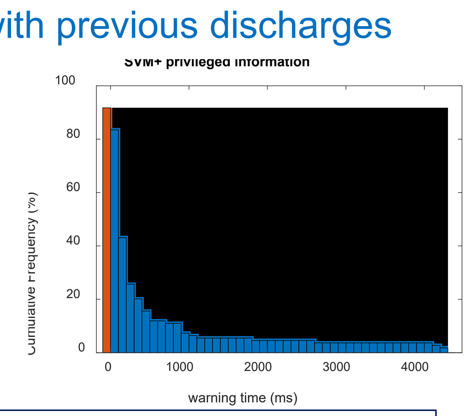


		Average warning time (ms)	Standard deviation (ms)
Success rate	**89.9%** (71/79)	-	-
Success rate with positive warning time	**82.3%** (65/79)	530	1020
Tardy alarm	**7.6%** (6/79)	3	1
False alarms	**5.9%** (23/392)	-	-

Overall statistics

- LID signal + mode lock as privileged information (JT-60SA situation)
- 801 non-disruptive discharges and 109 unintentional disruptions
- Initial predictor + 4 re-trainings
  - 9 missed alarms + 9 tardy detections but only 4 models with better results with previous discharges

		Average warning time (ms)	Standard deviation (ms)
Success rate	**91.7%** (100/109)	-	-
Success rate with positive warning time	**83.5%** (91/109)	380	880
Tardy alarm	**8.3%** (9/109)	3	2
False alarms	**3.2%** (26/801)	-	-



Conclusions

- Transfer learning has been tested with JET data: from C-wall to metallic wall
  - LID signal + mode lock as privileged information (JT-60SA present situation)
- The LUPI paradigm can be used for transfer learning with data scarcity
  - Adaptive predictors from scratch
  - It has been tested for transfer learning from C-wall to ILW (JT-60SA relevant for the future)
- Results are valid for mitigation purposes
  - On average, warning time O(100) ms and tardy alarms O(1) ms
- Future work
  - The LUPI paradigm has to be tested with more signals[7, 8]
  - Transfer learning from JET C-wall to JT-60SA C-wall
  - Transfer learning from JT-60SA C-wall to JT-60SA metallic wall (long time scale)

References

[1]J. . Vega et al. Nature Physics. 18, 741–750 (2022). DOI: 10.1038/s41567-022-01602-2
[2]C.G. Windsor et al. Nuclear Fusion 45, 5 (2005) 337. DOI: 10.1088/0029-5515/45/5/004
[3]A. Murari et al. Nuclear Fusion 60 (5) 056003
[4]A. Murari et al. Nuclear Fusion 61 (3) 036027
[5] V. Vapnik. The nature of statistical learning theory. Springer-Verlag
[6]V. Vapnik et al. Neural Networks 22 (2009) 544-557
[7]A. Murari et al. Nature Communications, **15, 2424 (2024). DOI: 10.1038/s41467-024-46242-7**
[8]R. Rossi et al. Plasma Physics and Controlled Fusion 64 (4) 045002. **DOI: 10.1088/1361-6587/ac4d3b**

Acknowledgements

This work has been partially funded by the *Spanish Ministry of Economy and Competitiveness* under the *Projects No. PID2022-137680OB-C31 and* PID2022-137680OB-C32

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
JT-60SA Advanced Superconducting Tokamak
BA-Satellite Tokamak Program
UNED
CONSORZIO RFX Ricerca Formazione Innovazione