Adaptive and Transfer Learning for Disruption Mitigation and Prevention on ASDEX-Upgrade and JET

Riccardo Rossi[1], Michela Gelfusa[1], Jesus Vega[2] and Andrea Murari[3], JET Contributors, ASDEX-Upgrade Team, and MST1 Team

1. Department of Industrial Engineering, University of Rome Tor Vergata, Via del Politecnico 1, 00133, Rome, Italy
2. Laboratorio Nacional de Fusion, CIEMAT, Avenida Complutense 40, Madrid 28040, Spain
3. Consorzio RFX (CNR, ENEA, INFN, Università di Padova, Acciaierie Venete SpA), Corso Stati Uniti 4, 35127 Padova, Italy

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Outline

1. **Introduction to the problem**
2. **Case study: from ASDEX Upgrade to JET***
   - **2.1 – Diagnostics**
   - **2.2 – Predictor**
   - **2.3 – Results on ASDEX-Upgrade**
   - **2.4 – Results on JET**
3. **Conclusions**

*Murari A., Rossi R., Peluso, E., Lungaroni M., Gaudio P., Gelfusa M., Ratta, G., Vega, J. and JET contributors "On the transfer of adaptive predictors between different devices for both mitigation and prevention of disruptions", Nuclear Fusion 2020

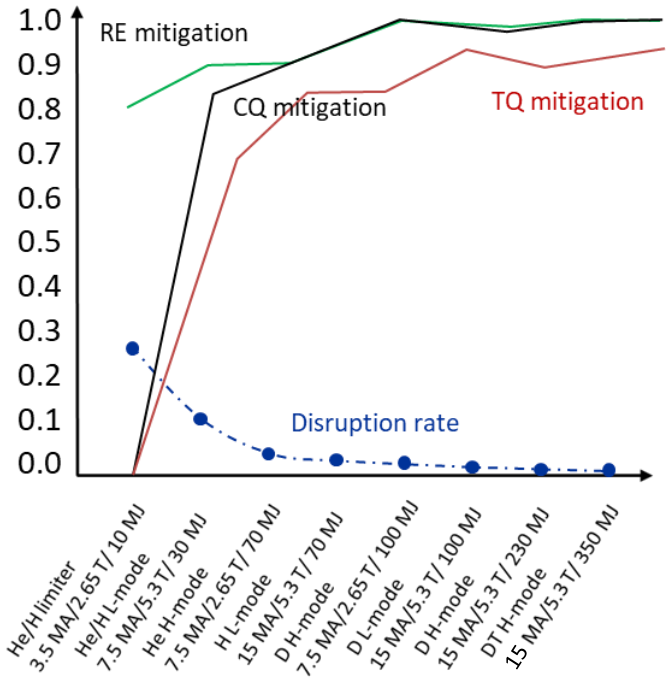
Introduction to the problem

Basic hypothesis of classic machine learning (i.i.d.) is violated in nuclear fusion.

How do we deal with:

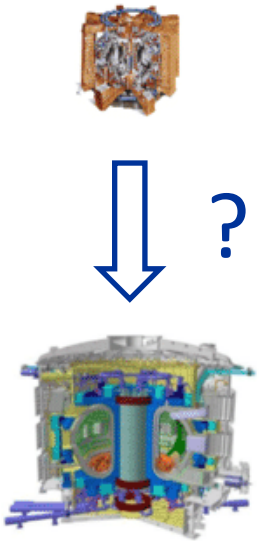
- Obsolescence?
- Transfer learning?

Obsolescence



M . Lehen “Challenges of Disruption Mitigation in ITER”, EPS 2017

Transfer learning



Possible solutions

Solution #1

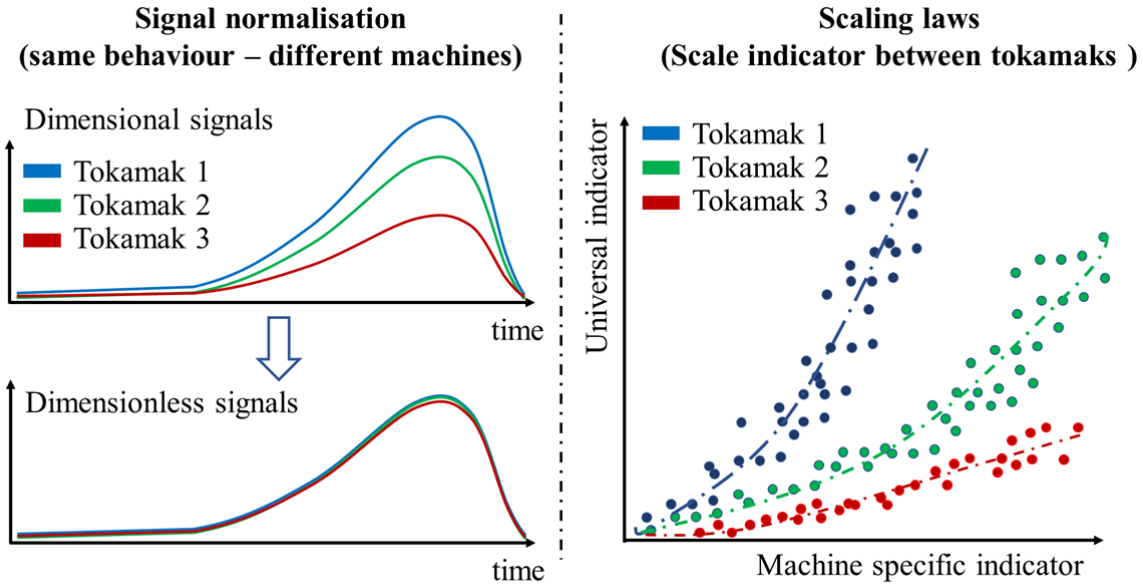
Rendering the training set and test set as “similar” as possible:

- **Dimensionless signals**
- **Scaling laws**

It is not always possible.

Extrapolation is needed for ITER, DEMO, etc.

Are extrapolation corrects? How large is the uncertainty?



P.C. de Vries et al 2016 Nucl. Fusion 56 026007

Possible solutions

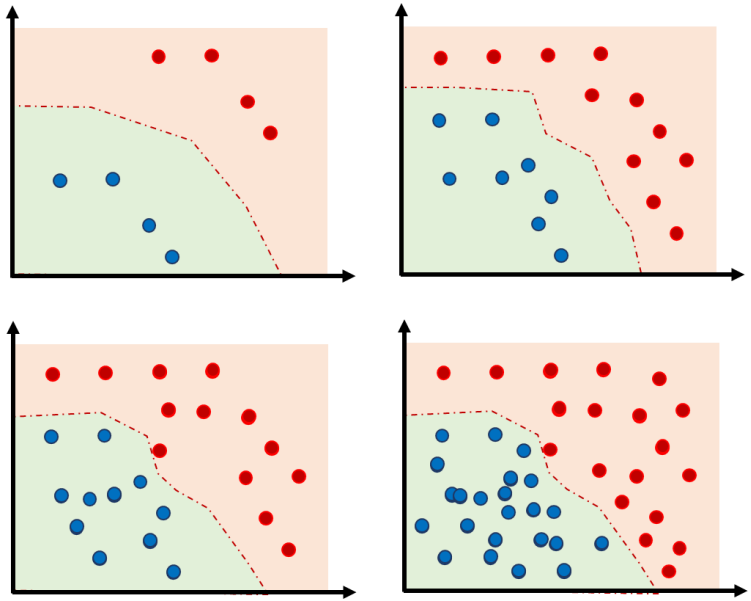
Solution #2

Updating the training set with the evolution experimental campaign:

- **Adaptive learning**

The training set is continuously updated (pulse after pulse) to achieve the best performances

Adaptive Machine Learning
(From scratch training, evolves after each pulse)



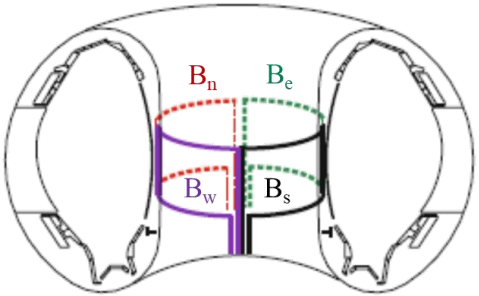
A.Murari et al "Prototype of an adaptive disruption predictor for JET based on fuzzy logic and regression trees" Nuclear Fusion 48(3) 2008
A. Murari et al "Adaptive predictors based on probabilistic SVM for real time disruption mitigation on JET" Nuclear Fusion, Volume 58, Number 5, March 2018

Case study from ASDEX-Upgrade to JET : Methodology

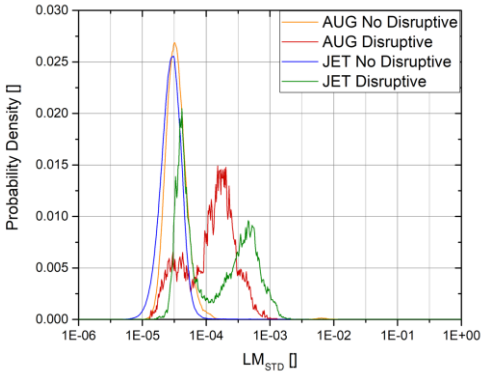
Diagnostics

1. Plasma current
2. Internal inductance
3. Locked mode
   Normalized with respect to the plasma current and tokamak geometry (similarly to *P.C. de Vries et al 2016 Nucl. Fusion 56 026007*)
4. Radiation (bolometers)
   Normalized indicators

Saddle coils for LM

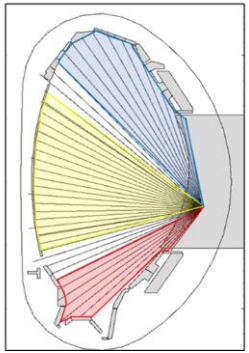


LM_{STD} pdf in AUG and JET

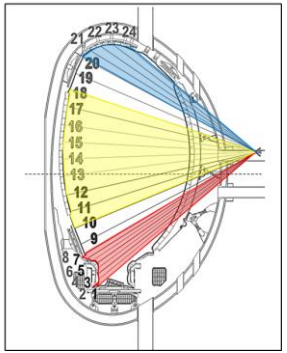


Bolometer Camera (Horizontal)

AUG Horizontal Camera



JET Horizontal Camera



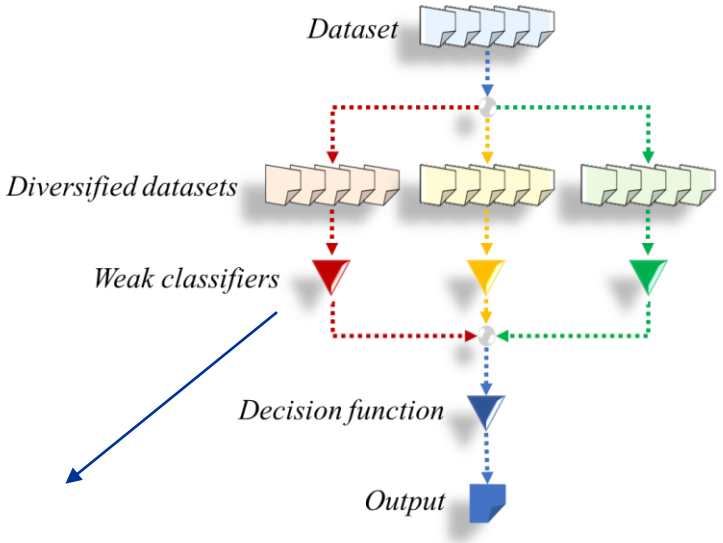
Case study from ASDEX-Upgrade to JET : Methodology

Predictor

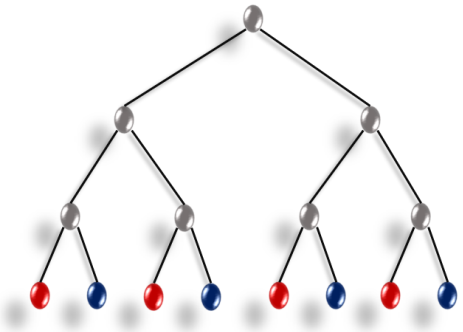
Ensemble of classifiers based on Classification and Regression Trees (**CART**)

Each classifier is trained with a diversified training set.

A decision function is used to provide a unique decision (disruptive or safe)

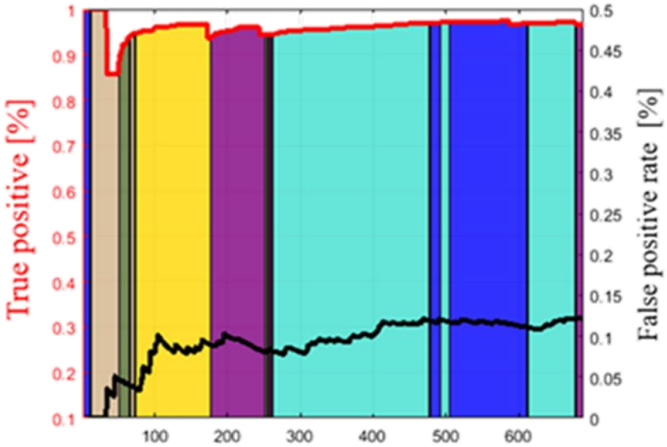


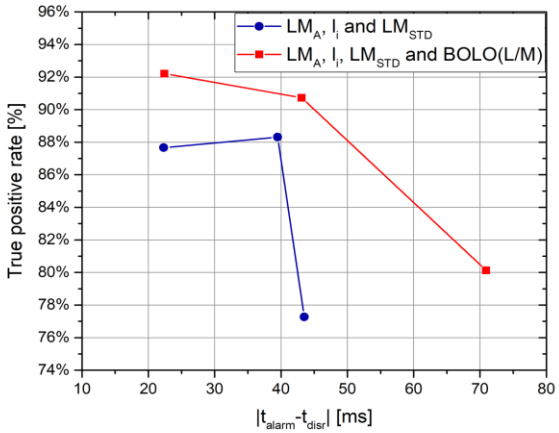
Trees per ensemble (Random Forests) = 40
Number of Ensembles = 11
Total weak classifier = 440

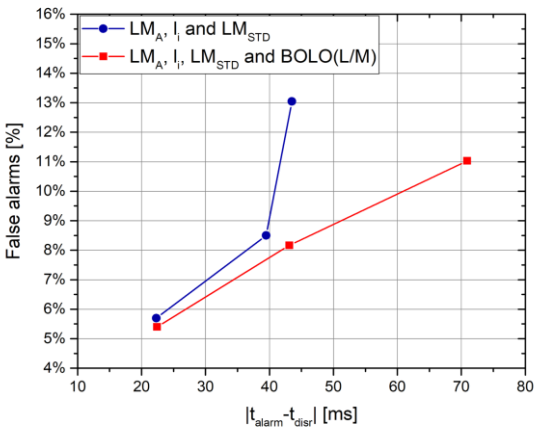


Case study from ASDEX-Upgrade to JET : Results

Results on ASDEX-Upgrade



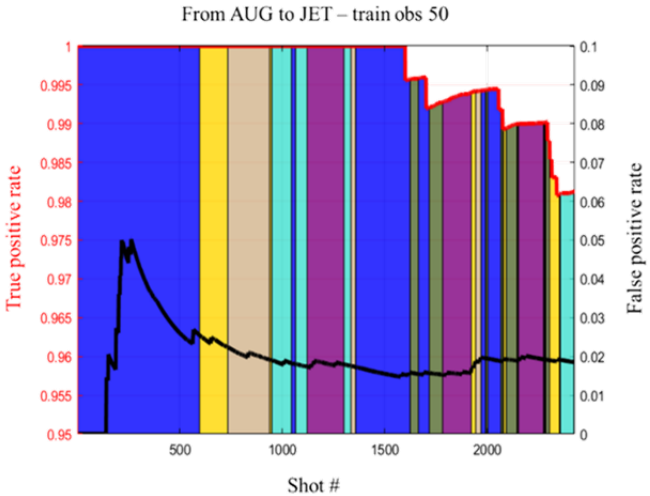


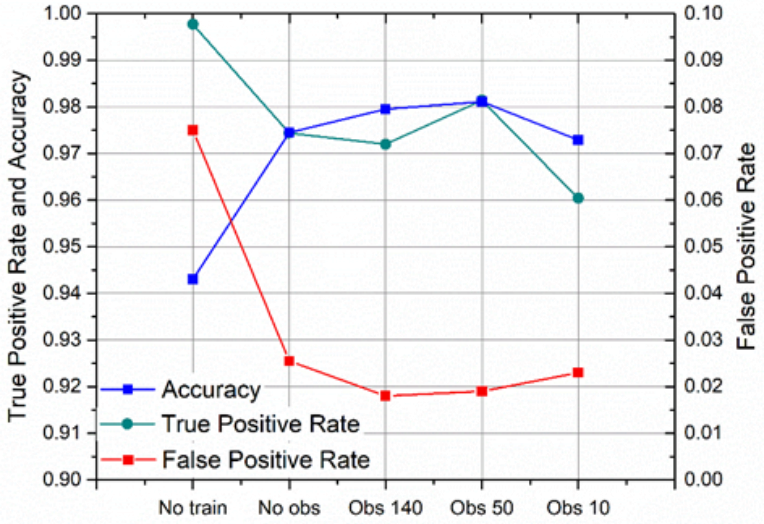


AUG	Success rate	Missed	Early	Tardy	False	Mean [ms]	Std [ms]
LM_A, li and LM_{std} 1.5 ms	87.66% (135/154)	5.84% (9/154)	5.84% (9/154)	0.65% (1/154)	5.70% (31/538)	22.3	66.1
LM_A L LM_{std} and Bolo L/M 10 ms	90.73% (137/154)	5.84% (9/154)	3.31% (5/154)	0.00% (0/154)	8.16% (44/539)	43.1	109.7

Case study from ASDEX-Upgrade to JET : Results

Results on JET without any specific training





JET	Success rate	Missed	Early	Tardy	False	Mean [ms]	Std [ms]
LM_A, l_i and LM_{std} 6 ms	98.14% (421/429)	1.40% (6/429)	0% (0/429)	0.47% (2/429)	1.90% (38/1998)	278.3	390.2
LM_A, l_i, LM_{std} and Bolo L/M 1000 ms	94.17% (404/429)	1.63% (7/429)	3.73% (16/429)	0.47% (2/429)	7.69% (150/1951)	489.7	664.9

Conclusions

Prediction on JET starting from a training on ASDEX-Upgrade has been tested and it has been demonstrated that acceptable prediction performances can be obtained if we use:

- Dimensionless signals;
- Adaptive approaches with transfer learning .

Moreover, predictors can be optimised to increase the warning time, but at the price of slightly lower accuracy.

Use of multiple predictors for different tasks (mitigation, prevention, and avoidance) may be a good solution, as also recently investigated by G. Rattà et al.

References

Murari A., Rossi R., Peluso, E., Lungaroni M., Gaudio P., Gelfusa M., Ratta, G., Vega, J. and JET contributors “On the transfer of adaptive predictors between different devices for both mitigation and prevention of disruptions”, Nuclear Fusion 2020

J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, A. Acero and JET-EFDA Contributors. “Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks”. Nuclear Fusion. 54 (2014) 123001 (17pp)

G. Pautasso et al. An adaptive real-time disruption predictor for ASDEX Upgrade. Nucl. Fusion 50 075004. 2010.

S. Dormido-Canto, J. Vega, J. M. Ramírez, A. Murari, R. Moreno, J. M. López, A. Pereira and JET-EFDA Contributors. “Development of an efficient real-time disruption predictor from scratch on JET and implications for ITER”. Nuclear Fusion. 53 (2013) 113001 (8pp).

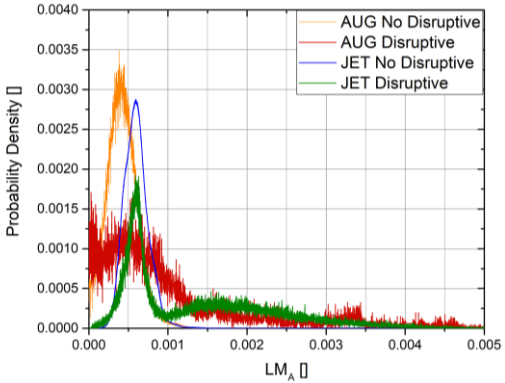
A. Murari et al “Adaptive predictors based on probabilistic SVM for real time disruption mitigation on JET” Nuclear Fusion, Volume 58, Number 5, March 2018

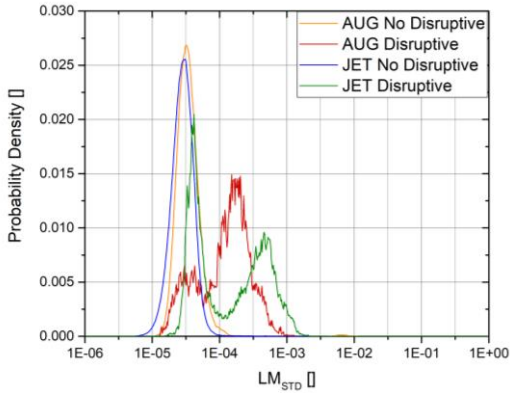
A.Murari et al “Prototype of an adaptive disruption predictor for JET based on fuzzy logic and regression trees” Nuclear Fusion 48(3) 2008

P.C. de Vries et al 2016 Nucl. Fusion 56 026007

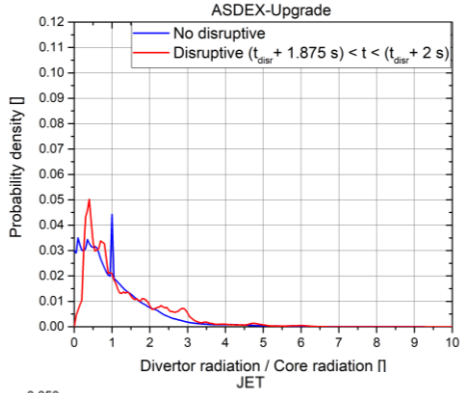
Appendix A – Data pdf

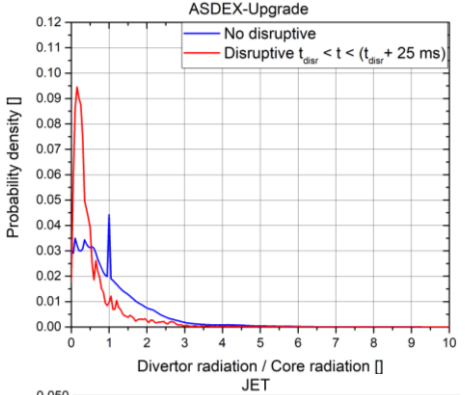
$$B_{LM,JET} = \sqrt{(B^{ns})^2 + (B^{ew})^2}$$
$$B_{LM,AUG} = \sqrt{(B^{ew})^2} = |B^{ew}|$$

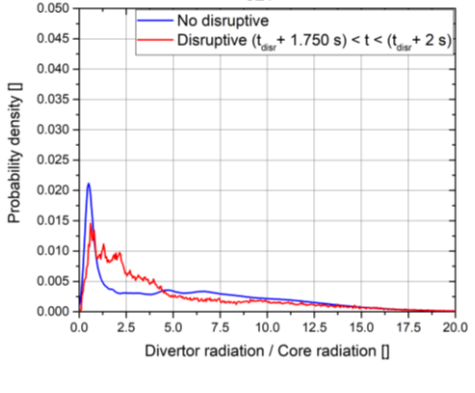


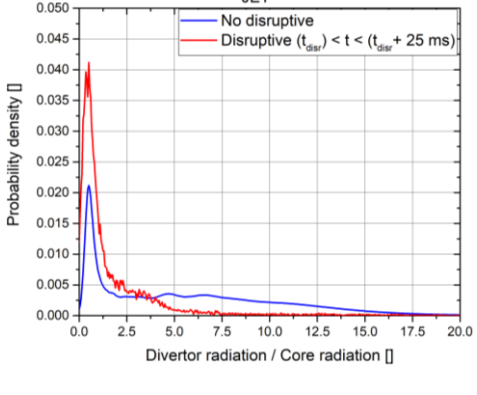


Divertor / Core Radiation Ratio







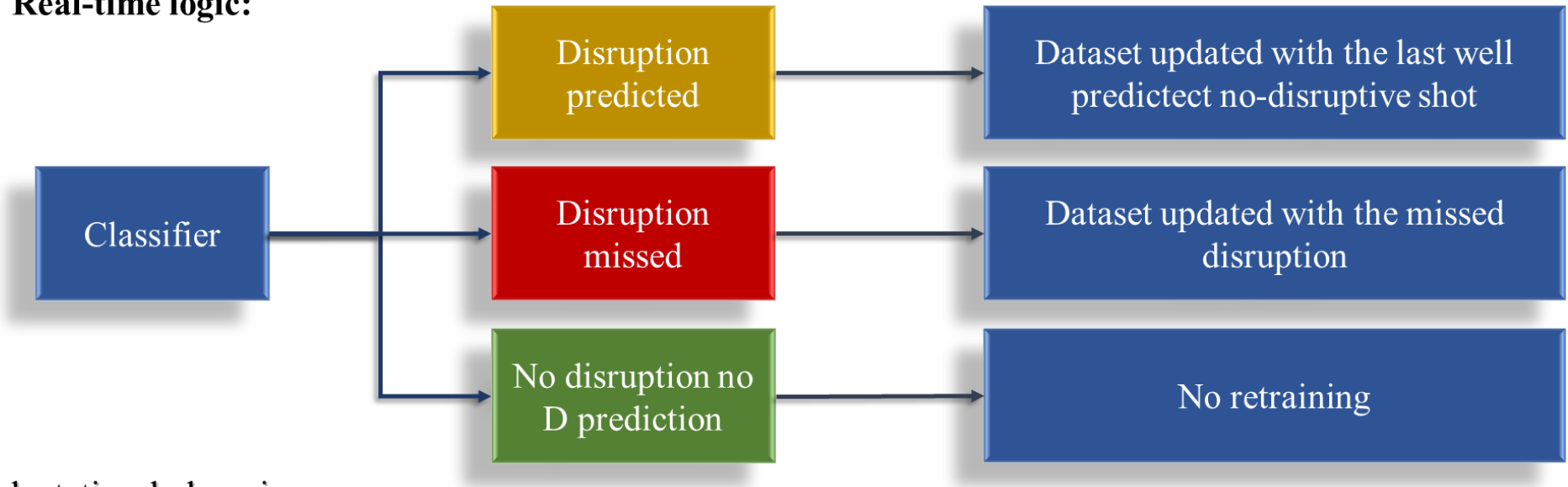


Appendix B – Decision functions and Obsolescence

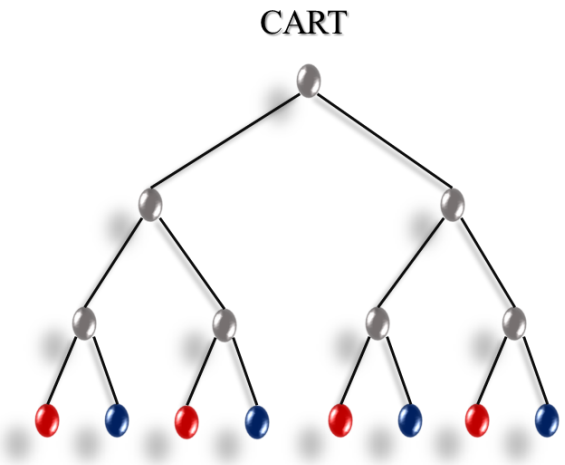
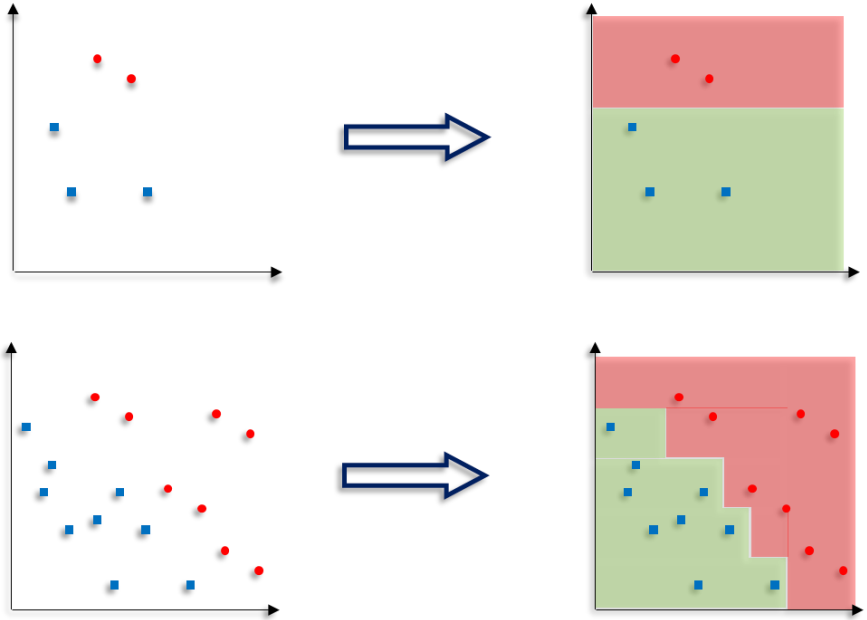
Adaptative learning

At each pulse, the dataset may be updated, and the classifiers retrained.

Real-time logic:



Adaptative de-learning

Oldest shots are removed from the datasets

Appendix C – Classification And Regression Tree (CART)