A Machine Learning Approach for Data Visualization and Parameter Selection for Efficient Disruption Prediction in Tokamaks

TH/7
ID: 1263

I. Bandyopadhyay[1,2,3,a] , Y. K. Meghrajani[4], S. Patel[5], J. Patel[5], H. S. Mazumdar[5], L. S. Desai[6], V. K. Panchal[2,3], R. L. Tanna[2,3], J. Ghosh[2,3] and the ADITYA Team[2]

[1]ITER-India, Institute for Plasma Research, Bhat, Gandhinagar, India
[2]Institute for Plasma Research, Bhat, Gandhinagar, India
[3]Homi Bhabha National Institute (HBNI), Mumbai, India
[4]Department of Electronics & Communication, Faculty of Technology, Dharmsinh Desai University, Nadiad, India
[5]Research & Development Center, Dharmsinh Desai University, Nadiad, India
[6]Department of Mathematics, Faculty of Technology, Dharmsinh Desai University, Nadiad, India
[a]Email: indranil.bandyopadhyay@iter-india.org

Abstract: Disruption prediction in tokamaks is a challenging task and often involves processing of huge amount of diagnostic database, either through physics based or artificial intelligence based detection models. This paper presents a novel tool for quick data visualization and parameter selection for disruption prediction based on a machine learning technique using ADITYA tokamak data. This study involves a data set of 2216 ADITYA discharges, including both disrupting and non-disrupting shots. Firstly, a subset of 1000 labeled shots is utilized, each having 156 recorded parameters, thus a total dataset of 156000 data sets. An elaborate offline artificial neural network (ANN)-based correlation algorithm is developed to compute the score of each measured diagnostic parameter with respect to plasma current. The combined result of the ANNs presently predicts the shot-type with overall 97.11% accuracy, whereas share of disruption classification accuracy is 99.0%. This is possibly due to the subset of database with human errors, which is taken care of by the neural network. Elaborate visualization tool with 2D color-coded map for the entire dataset is developed for easy decision making. Furthermore, this tool will be useful to develop a numerical system for prediction of plasma disruptions using state-of-the-art machine learning algorithms applied on diagnostic data, which will be compatible with the real time hardware-based solution for avoidance or mitigation actions.

Motivation

- For tokamaks to be attractive as the core of future fusion based power plants, it must operate in steady state or at least quasi-steady state without plasma current disruptions.
- Early and effective disruption prediction is essential for disruption avoidance or mitigation.
- Disruptions are multi-dimensional in nature. Physics based prediction is difficult as it needs simultaneous tracking of many plasma and machine parameters for effective disruption prediction.
- Perhaps a better way for effective predictions technique would be based on machine learning technique, which has received a lot of interest in recent times across several tokamaks.
- ADITYA is a small tokamak, but excellent for disruption studies.
- Tools based on Machine learning algorithms are developed for efficient ADITYA data visualization, parameter selection for disruption analysis and finally disruption detection and prediction
- Final goal is to predict disruption events 16-20 msec prior to disruptions in ADITYA with >99% accuracy

1

ADITYA Database used for this work

- Total 2216 plasma discharge data of 14000, 25000 and 26000 series are used including both disrupting and non-disrupting shots
- 14000 series was already used earlier by Sharma et al for disrution prediction using a convoluted neural net
- Each of the diagnostic parameter of these shots have 2048 samples collected at a sampling rate of 5 kHz

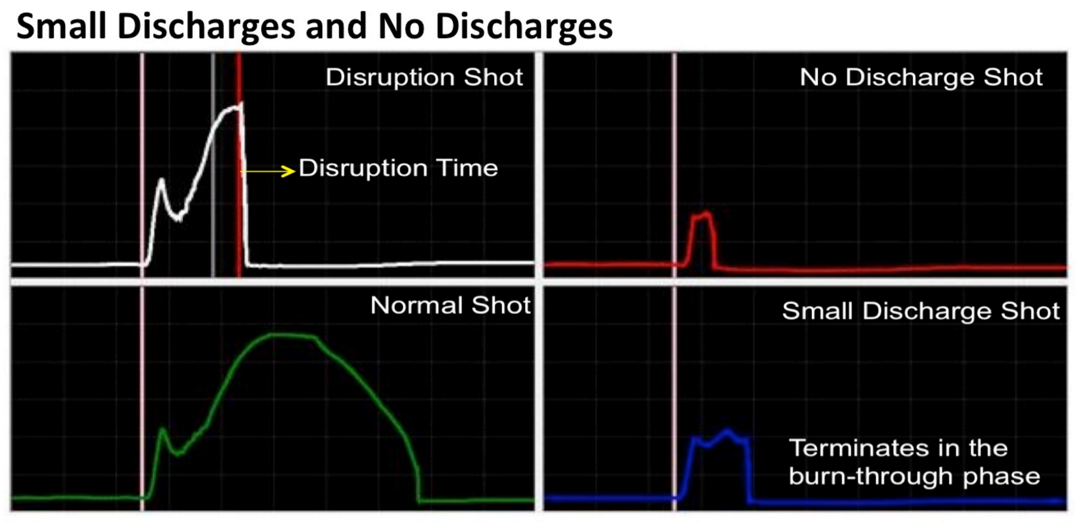
Ser.	Para.	Shots Used	Year	Label			Sampling Rate (kHz)		Data Range (msec)	
				Nor.	Dis.	Oth.	Min	Max	Min	Max
14000	96	216	2004	80	110	36	5	5	-51	358.4
25000	181	1000	2012	354	347	299	5	1000	-409.4	409.6
26000	214	1000	2013	0	0	0	5	1000	-409.4	409.6

Ser=ADITYA Data Series; Para=Number of Parameters; Nor=Normal; Dis=Disrupted; Oth= Other type

2

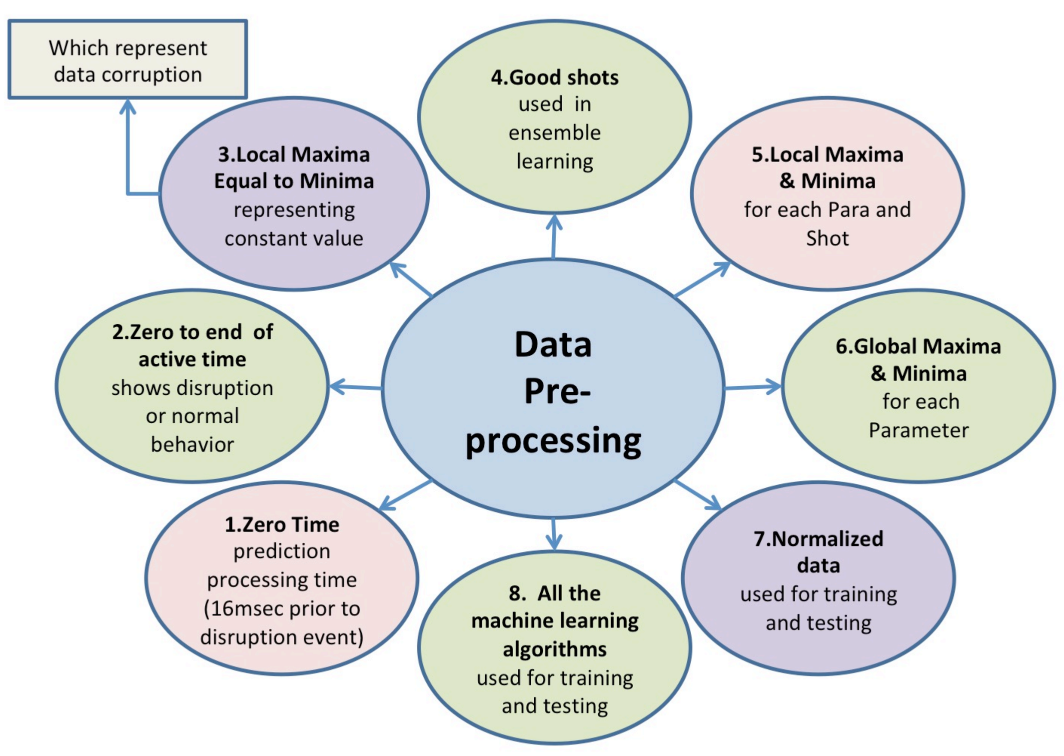
Data Viewer & Automated Shot Classification

- ADITYA tokamak diagnostic data viewer is designed using C#.NET framework for rapidly examining 20736 time series dataset (216 x 96) consisting of approximately 50 million data points.
- ANN tools developed for automated classification of the discharges inFour shot types – Normal shots, Disruption shots, Small Discharges and No Discharges



3

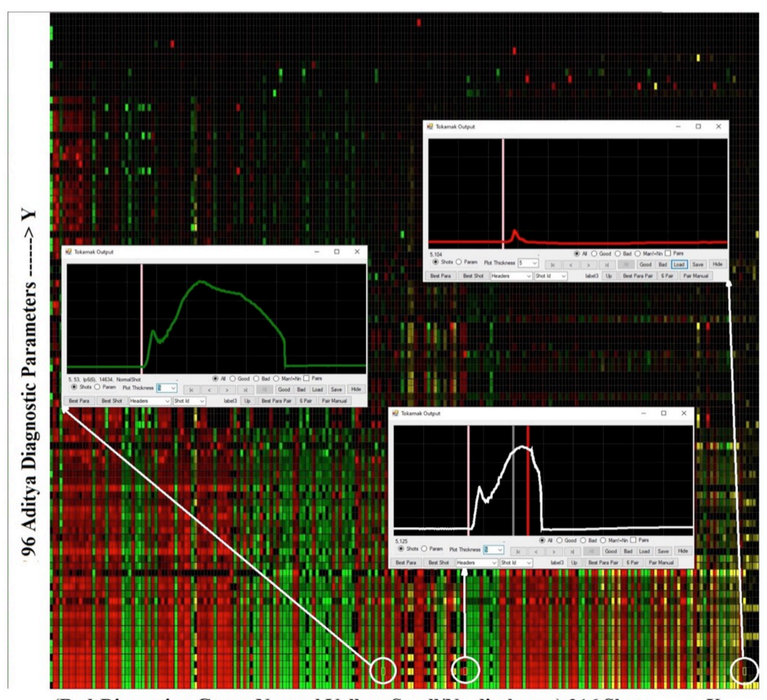
Data Selection and Disruption Prediction using ANN Tools



4

2D Color-Coded Entropy Plots

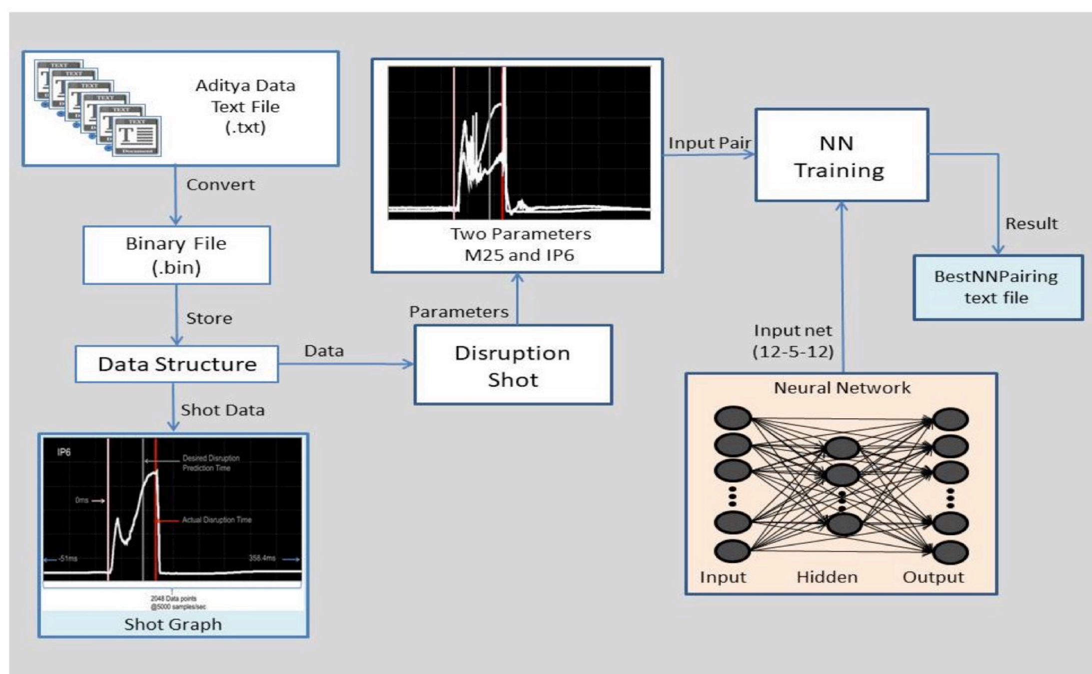
- Color coded Entropy Plot - single panel contains data of 216 shots x 96 parameters for each shot
- Information density is highest to lowest across the diagonal, with top right corner with minimum information and bottom left corner with maximum information



(Red:Disruption Green:Normal Yellow:Small/No discharge) 216 Shots------> X

5

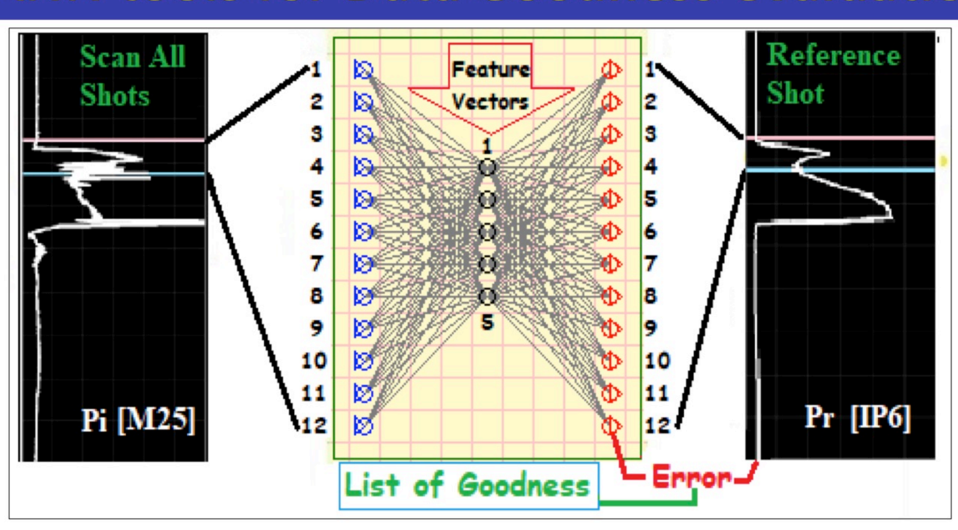
Data Goodness Classification using ANN



- Vector Quantize Code Book for Data Goodness Classification
- 15-45 msec data is used for all the diagnostic parameters

6

ANN tools for Data *Goodness* evaluation



- A 3-layer ANN is used to extract and encode the feature vectors of 12 input points of parameter of the shot in its hidden 5 nodes. These 5 feature vectors are decoded to reconstruct 12 output points of reference parameter of the shots
- The Net is trained for 1000 cycles. The inverse of error between Net output and Pr represents the goodness of feature match. All 96 Pi parameters are sorted according to their goodness numbers as parameter selection list to be used for shot type prediction.

7

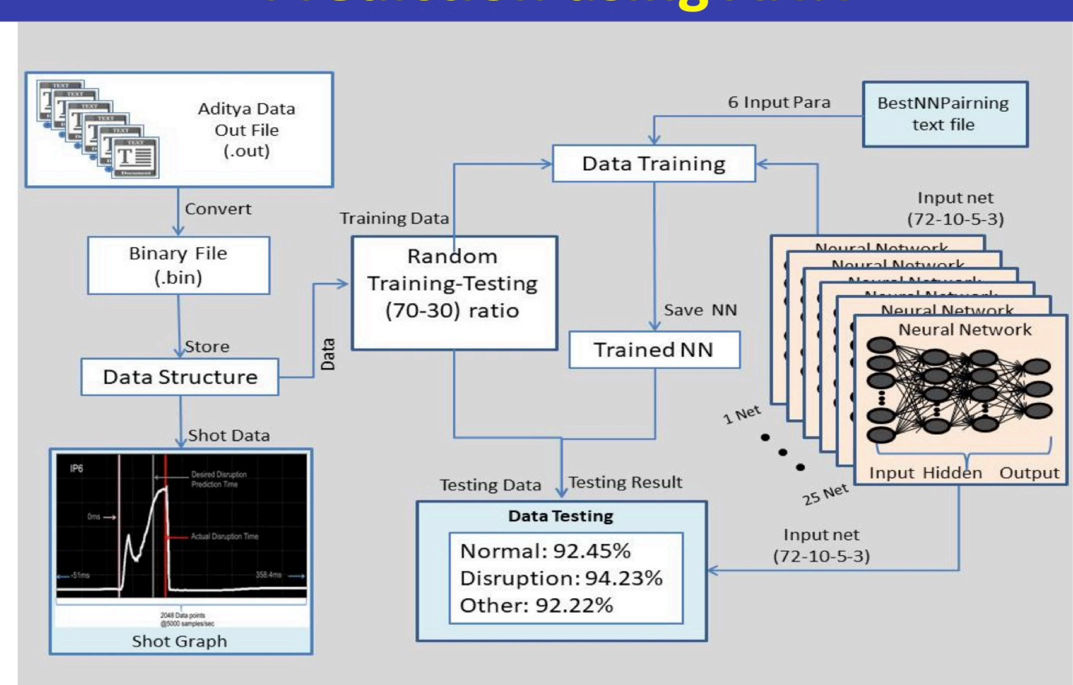
Data Goodness Score

Sr. No.	Top 12 Parameters Name in sorted order (channel no)	Parameter description	Goodness Score(%)of Parameter with respect to Plasma Current
1	Ip5(5)	Rogowski Coil	97.68
2	B1(65)	Radial bolometry	92.61
3	GIM(30)	Grazing Incidence Monochromator	92.56
4	M3(10)	Mirnov Coil	92.02
5	B5(69)	Radial bolometry	86.02
6	B3(67)	Radial bolometry	85.93
7	B6(70)	Radial bolometry	85.59
8	BOLO1(17)	Top bolometry	84.11
9	B11(75)	Radial bolometry	83.82
10	M4(11)	Mirnov Coil	82.85
11	B10(74)	Radial bolometry	81.58
12	M1(8)	Mirnov Coil	81.54

All correlations are measured with respect to IP6 Rogowski coil measurements for plasma current

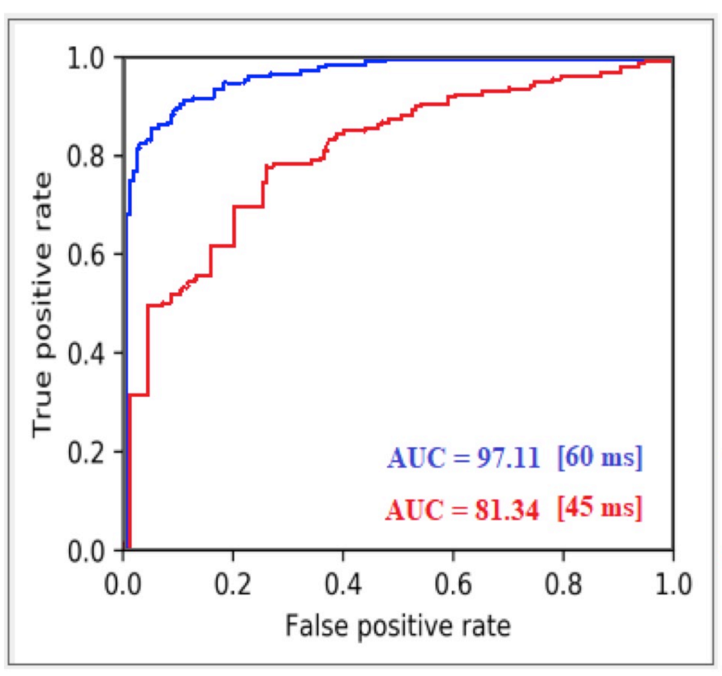
8

Ensemble Learning and Disruption Prediction using ANN



9

Disruption Learning and Prediction using ANN



- 2000 shots for 25000 and 26000 series are randomly split in 70-30 ratio for training & testing
- 0-60 msec data for training and 0-45 msec data used for testing
- Blue curve Learning curve
- Red curve prediction of disruptive shots

10

Summary

- Novel tools for quick data visualization and parameter selection for disruption prediction based on a machine learning technique using ADITYA tokamak data have been developed.
- This study involves a data set of 2216 ADITYA discharges with 1D time series data, including both disrupting and non-disrupting discharges.
- Final goal is to predict disruption events 16-20 msec prior to disruptions in ADITYA with >99% accuracy
- The combined result of the ANNs presently predicts the shot-type with overall 97.11% accuracy, whereas share of disruption classification accuracy is 99.0%.
- Prediction accuracies will be further improved by inclusion of 2D profile database.
- These tools will further be applied on ITPA multi-machine disruption database

11

References

1. A. Sengupta and P. Ranjan 2000 Nucl. Fusion 40 1993.
2. A. Sengupta and P. Ranjan 2001 Nucl. Fusion 41 487.
3. C.G. Windsor *et al.* 2005 Nucl. Fusion 45 337.
4. M. Cacciola, *et al.* 2006, International Journal of Intelligent Technology, 1.4, pp. 274-280.
5. A. Murari *et al.* 2009 *Nucl. Fusion* 49 055028.
6. William Tang *et al.* 2016, 26th IAEA Fusion Energy conference, 17-22 October, Kyoto, Japan, EX/P6-47.
7. Matthew S Parsons 2017 Plasma Phys. Control. Fusion 59 085001.
8. C Rea *et al.* 2018 *Plasma Phys. Control. Fusion* 60 084004.
9. A. Murari *et al.* 2018 Nucl. Fusion 58 056002.
10. W. Zheng *et al.* 2018 Nucl. Fusion 58 056016.
11. P. Sharma *et al.* 2018, Towards Extensible and Adaptable Methods in Computing, pp. 179-193.
12. K.J. Montes et al 2019 Nucl. Fusion 59 096015.
13. Julian Kates-Harbeck *et al.* 2019, Nature, 568 (7753).
14. Y. Fu *et al.* 2020, Physics of Plasmas 27.2 (022501).
15. A. Mehta and H. Mazumdar 2020, Computational Biology and Chemistry, 84 (107164).
16. M. Patel and H. Mazumdar 2014, Journal of theoretical biology, 361, pp. 182– 189.

12

28th IAEA Fusion Energy Conference (FEC2020), 10-15 May 2021 *(Virtual Event)*

ITER-India
प्लाज़्मा अनुसंधान संस्थान
Institute for Plasma Research
DHARMSINH DESAI UNIVERSITY