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

I. Bandyopadhyay[1,2,3] 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

Outline

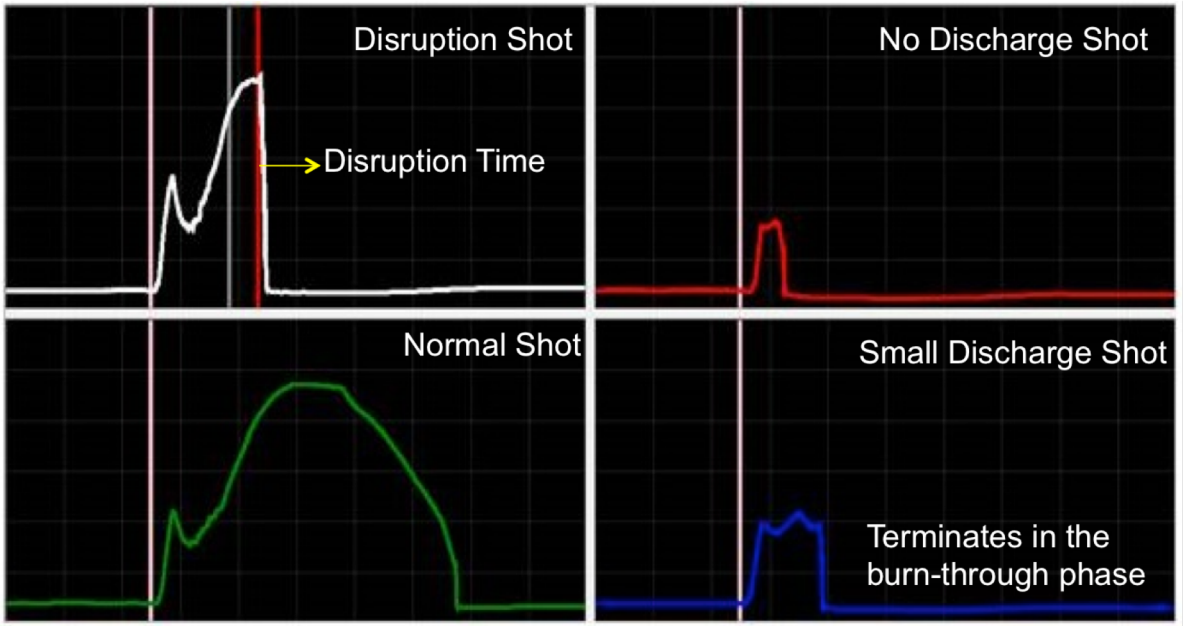
- ADITYA database used for analysis using artificial neural net (ANN)
- Application of ANN for
  - Automatic Classification of ADITYA shots
  - Data Entropy Scope for 2D visualization of large database
  - Data goodness classification for use in ANN based prediction model
  - Disruption learning and prediction
- Summary and Future Work

ADITYA database used

- Total 2216 plasma discharge data of 14000, 25000 and 26000 series are used including both disrupting and non-disrupting shots
- ANN tools developed for automated classification of the discharges
- Four shot types – Normal shots, Disruption shots, Small Discharges and No Discharges
- Total 2048 samples per shot @ sampling rate of 5000 samples/second.

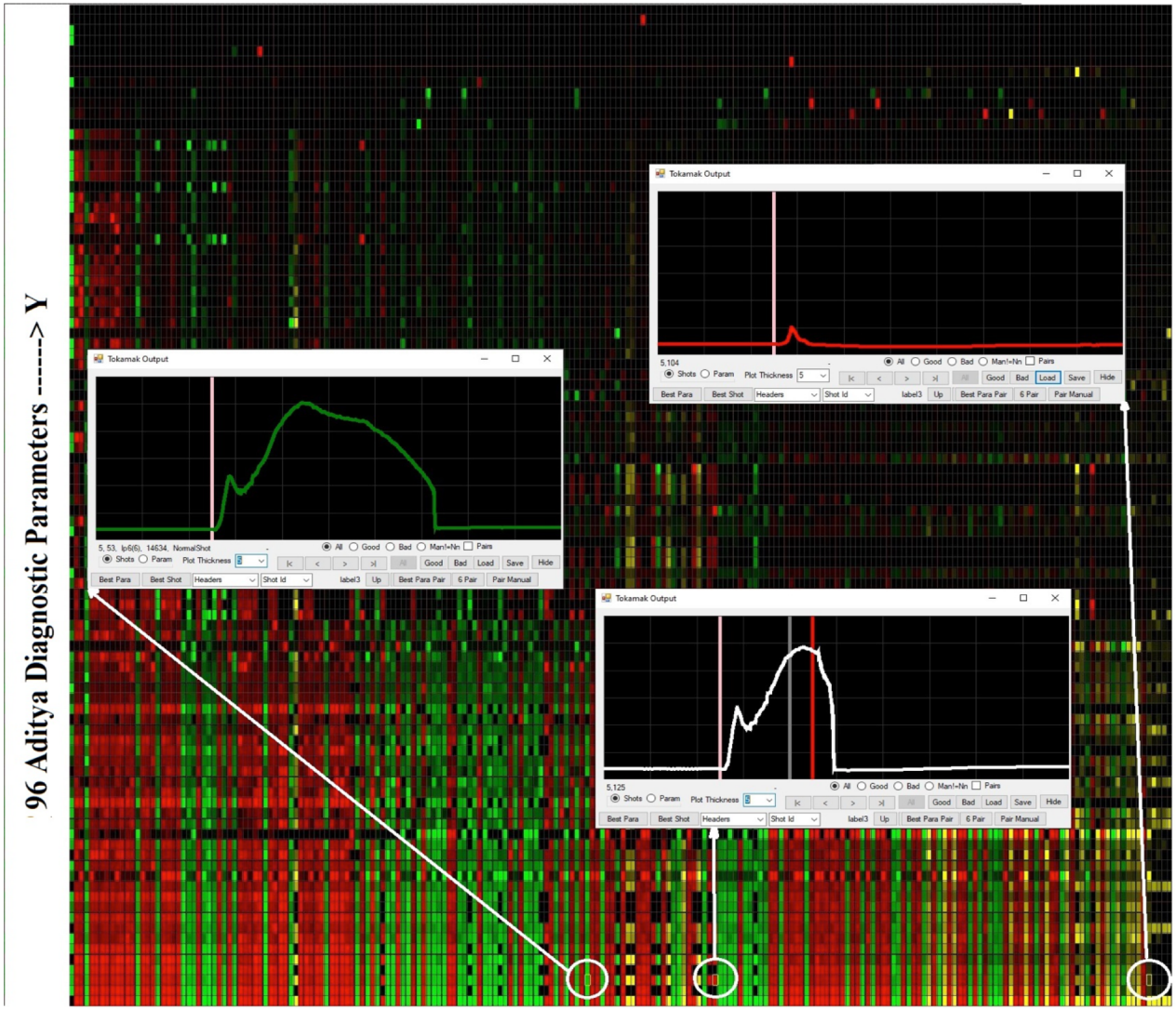
Ser.	Para.	Shots Used	Year	Label			Sampling Rate (kHz)		Data Range (msec)	
				Nor.	Dis.	Oth.	Min	Max	Min	Max
14000	96	216	2004	80	110	36	5	5	-51	358.4
25000	181	1000	2012	354	347	299	5	1000	-409.4	409.6
26000	214	1000	2013	0	0	0	5	1000	-409.4	409.6

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



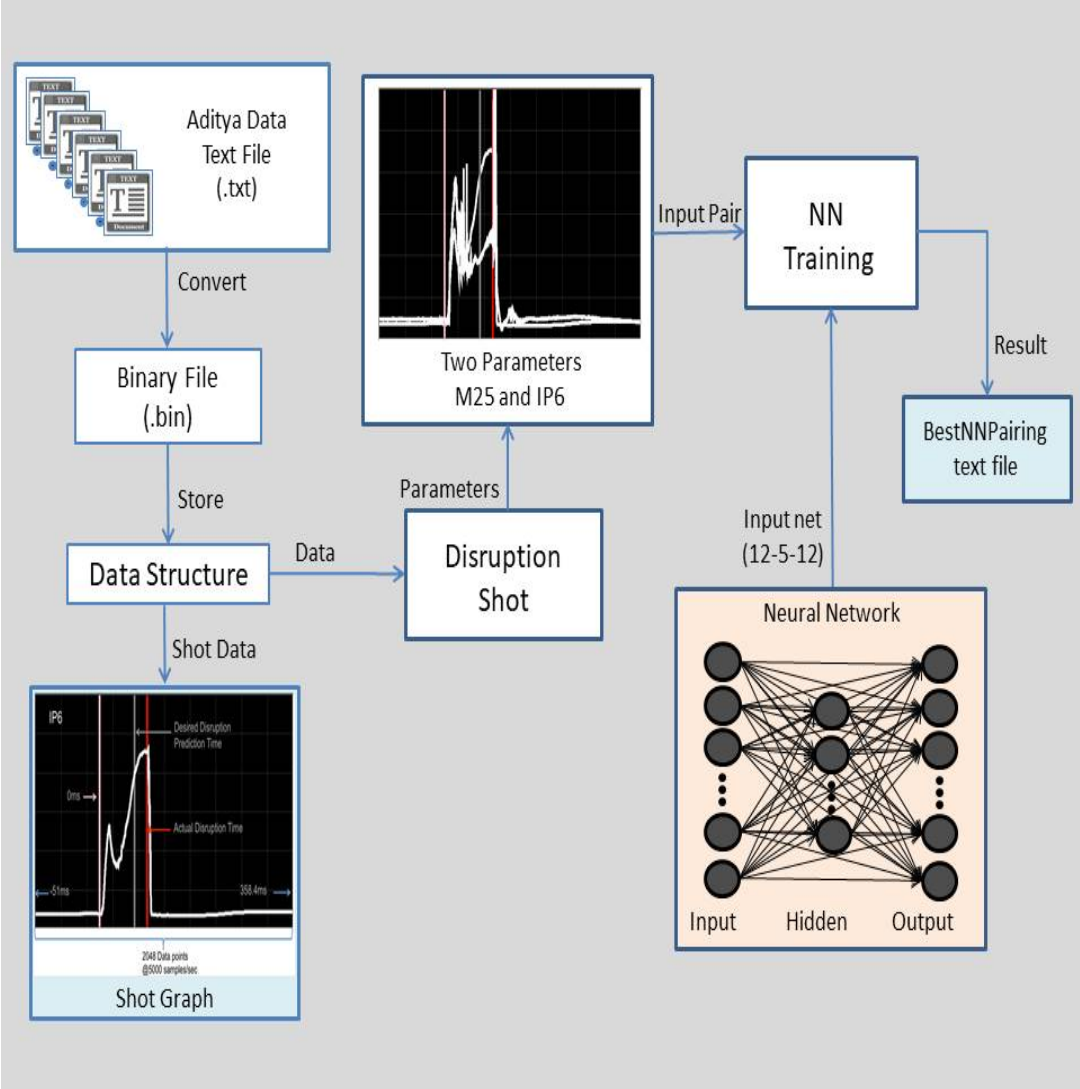
2D Data Entropy Plots

- 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



Data Goodness Classification using ANN

- Vector Quantize Code Book for Data Goodness Classification
- A 3-layer ANN to extract and encode the feature vectors of 12 input parameter, denoted as Pi, in its hidden 5 nodes, which are decoded to reconstruct 12 output parameter denoted as Pr.
- The Net is trained for 1000 epochs.
- 15-45 msec data is used for all the diagnostic parameters

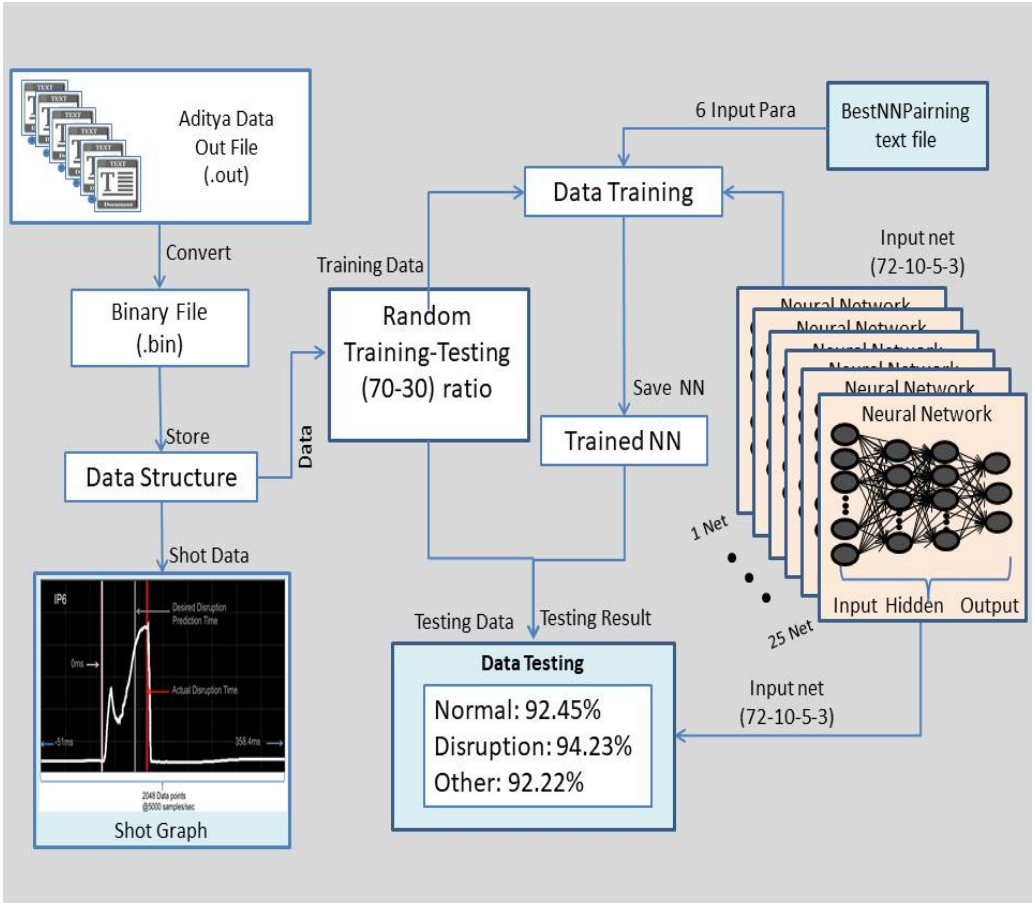


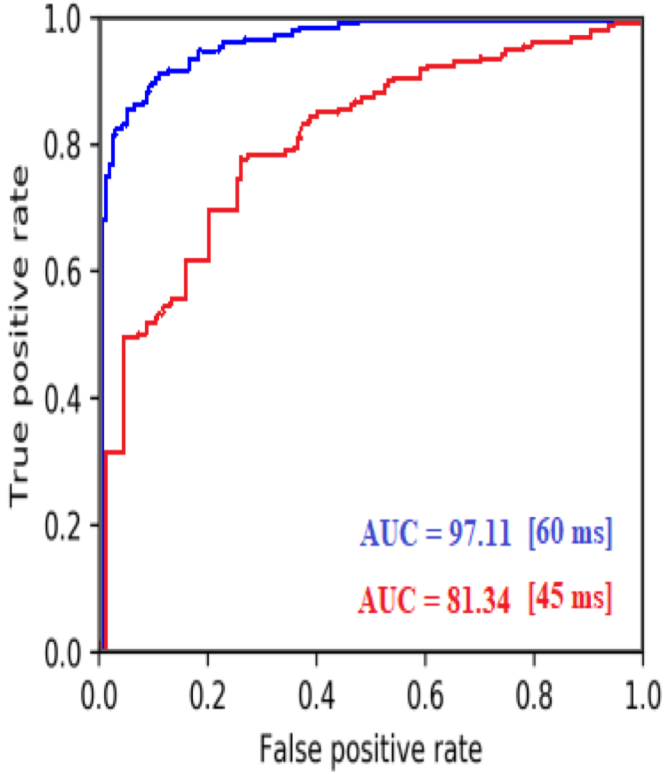
Data Goodness Score

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

Disruption 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

Summary and Future Work

- A novel tool for quick data visualization and parameter selection for disruption prediction based on a machine learning technique using ADITYA tokamak data has 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 16-20msec 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