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

I. Bandyopadhyay1,2,3, Y. K. Meghrajani1, S. Patel3, J. Patel3, H. S. Mazumdar2, L. S. Desai6, V. K. Panchal2,3, R. L. Tanna2,3, J. Ghosh2,3 and the ADITYA Team2

1ITER-India, Institute for Plasma Research, Bhat, Gandhinagar, India
2Institute for Plasma Research, Bhat, Gandhinagar, India
3Homi Bhabha National Institute (HBNI), Mumbai, India
4Department of Electronics & Communication, Faculty of Technology, Dharmsinh Desai University, Nadiad, India
5Research & Development Center, Dharmsinh Desai University, Nadiad, India
6Department of Mathematics, Faculty of Technology, Dharmsinh Desai University, Nadiad, India

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 ones. 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 cannot be predicted in dimensional nature. Physics based prediction is difficult as it requires 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.

ADIITYA 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 disruption prediction using a novel neural net.
- Each of the diagnostic parameter of these shots have 2048 samples collected at a sampling rate of 5 kHz.

Data Viewer & Automated Shot Classification

- ADITYA tokamak diagnostic data viewer is designed using CR.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 disruptions into shoot types – Normal shots, Disruption shots, Small Discharges and No Discharges.

Data Selection and Disruption Prediction using ANN Tools

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

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.

Data Goodness Score

- All correlations are measured with respect to IPS Rogowski coil measurements for plasma current.

Ensemble Learning and Disruption Prediction using ANN

- Ensemble of ANNs are used for disruption prediction, which are trained on data from each other.
- Datasets are divided into training and testing sets.
- Ensemble of ANNs are used for disruption prediction, which are trained on data from each other.

Summary

- Novel tools for quick data visualization and parameter selection for disruption prediction based on 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%.
- These tools will further be applied on ITPA multi-machine disruption database.

References