MACHINE LEARNING MODELS FOR REAL-TIME INFERENCE OF PLASMA DYNAMICS USING THE 2-DIMENSIONAL BEAM EMISSION SPECTROSCOPY SYSTEM AT DIII-D

Classification, onset prediction and forecasting of edge localized modes and confinement regimes

L. MALHOTRA[1]*, D.R. SMITH[1], P. ARORA[1], G. MCKEE[1], Z. YAN[1], J. ZIMMERMAN[1], M.D. BOYER[2], R. COFFEE[3], A. JALALVAND[4], E. KOLEMEN[5]

[1] University of Wisconsin-Madison, Madison, WI, USA

[2] Princeton Plasma Physics Lab, Princeton, NJ, USA

[3] SLAC National Accelerator Lab, Menlo Park, CA, USA

[4] Ghent University, Ghent, Belgium

[5] Princeton University, Princeton, NJ, USA

*Email: lmalhotra@wisc.edu

Abstract

Multi-channel fluctuation diagnostics capture the spatial patterns of high-bandwidth plasma dynamics. The paper reports on an effort to develop machine learning (ML) models for the real-time identification of edge-localized-mode (ELM) events and the turbulence properties of confinement regimes using the 2-dimensional Beam Emission Spectroscopy (BES) system at DIII-D. The 64-channel BES system captures plasma density perturbations at a 1 MHz frame rate. The "edge ML" models will be deployed on high-throughput hardware such as a Field Programmable Gate Array (FPGA) accelerator for integration in the real-time plasma control system (PCS). The models will generate reduced signals that correspond to ELM activity and turbulence dynamics, and the real-time PCS will learn to avoid ELM regimes and to achieve and sustain advanced confinement regimes such as the wide pedestal QH-mode.

The ELM onset dataset contains about 400 hand-labelled ELM events each consisting of about 4-8 millisecond time series data from all the 64 BES channels. Before model training, about 15% of the total ELM events are kept for cross-validation and 10% of them are reserved as previously unseen (test) data. The remaining data is used for training various machine learning models ranging from logistic regression to deep convolutional neural networks. The entire training and cross-validation workflow makes use of time series data truncated beyond 75 µs after the leading edge of the active ELM in each ELM event. Instead of sequential training on the time series, the input data windows, and their corresponding labels, are sampled at random from the ELM events. This whole data pre-processing and training pipeline is repeated for different values of label lookaheads from 10-100 µs to assess the predictive power of the trained models in context of predicting ELM onset and turbulence activity. The model performance on the test data is analyzed using metrics like precision, recall, F1-score and area under the receiver operating characteristics curve (ROC curve) showing the variation of True Positive Rate (TPR) with False Positive Rate (FPR). The model robustness is also assessed using custom metrics such as macro predictions in which the entire time series is divided into the regions of pre-active ELM (region on the left of the first vertical line in the left figure), and active ELM (region between two vertical lines) and model predictions are analyzed in each of the regions separately which are then combined. Various feature engineering techniques based on the max and average pooling or the output feature map from a convolutional neural network are also employed and tested against classical machine learning algorithms such as logistic regression, random forests and gradient boosting. As compared to the baseline results of around 0.83 using classical machine learning algorithms, deep convolutional neural networks yield an AUC-ROC score of about 0.92 on the test data (right figure). Preliminary results also indicate that deep convolutional models can predict ELM onset up to 100 µs in advance using data windows as small as 16 µs. The predictive performance suggests that ELM precursors can be identified and detected in high-bandwidth data streams for ELM onset warning. The paper also explores other data pre-processing techniques like spatial and time gradients as color channels and different neural network architectures such as autoencoders to perform multiple classification and forecasting tasks like ELM prediction, LH transition forecasting, and control for enhanced confinement regimes such as the wide pedestal QH-mode. Autoencoders can help to compress the spatial-temporal information from the input data windows in a low-dimension feature space which can be used for self-supervised feature engineering with unlabelled datasets.