ARTICLE IN PRESS

Fusion Engineering and Design xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Fusion Engineering and Design

journal homepage: www.elsevier.com/locate/fusengdes

Automatic detection of L-H transition in KSTAR by support vector machine

Gi Wook Shin[a], J.-W. Juhn[b], G.I. Kwon[b], S.H. Son[b], S.H. Hahn[a,b,*]

[a] Korea University of Science and Technology, 217 Gajeong-ro, Yuseong-gu, 34113, Republic of Korea
[b] National Fusion Research Institute, 113 Gwahak-ro, Yuseong-gu, 34133, Republic of Korea

ARTICLE INFO

Keywords:
KSTAR
Plasma control
Machine learning
L-H transition
Support vector machine

ABSTRACT

Method for automatic detection of L-H transition using Support Vector Machine (SVM), a popular tool of supervised machine learning tools, has been evaluated in order to improve plasma density control in KSTAR. Through the SVM, a nonlinear classifier is trained to distinguish L-mode and H-mode using two kinds of diagnostic data measured in KSTAR. The trained classifier has been analyzed for possible usage on the real-time detection through the truncation of the training samples. Study on the optimization of the training samples, and corresponding accuracy change is made for evaluating feasibility for real-time implementations.

1. Introduction

In many tokamaks including KSTAR [1], a high confinement mode (H-mode) [2,3] operation is highly desired for more efficient fusion plant design. The H-mode plasma has higher density pedestal than a low confinement mode(L-mode) on the plasma edge hence it shows better performances than the L-mode. It is widely known that the density control efficiency is very different for H-mode plasmas, mainly due to the change of fueling efficiency related to the edge transport barrier at the H-mode. Therefore, it is generally recommended to use different fueling devices, different diagnostics and even different algorithms. In order to apply different algorithm for L- and H-mode plasmas, it is required to determine whether the plasma is in the L-mode or H-mode and choose the appropriate control scheme according to the state of the plasma.

In this work, we did a feasibility study on automatic detection of the L-H transition event during a KSTAR plasma pulse by utilizing the method of Support Vector Machine (SVM) [4]. As seen from previous studies [5–7] using SVM, classifier using the SVM shows high classification performance. Method of data selections, performance check and optimizations for the real-time applications are described. In order to classify the L-H transition properly in real-time, we aim to obtain a detection probability above 80%. From L-H transition to the first ELM burst, it takes several tens of milliseconds to a few hundred milliseconds. Therefore, the faster the calculation time for classification, the more detections our classifier can have. Considering a classifier within 10–20 ms computational cycle and above 80% accuracy, we can have enough detection opportunities after the L-H transition and then we can classify whether plasma is L- or H-mode before first ELM burst.

Accuracy and performances on the models consisting a single feature and double features are evaluated. Finally, an appropriate candidate of the model constructed using an optimized set of samples is suggested as a result.

2. Method

The SVM is one of the tools of supervised machine learning, which is a classifier that has the advantage of having no local minimum problem because it learns by minimizing a convex function. In this chapter, we describe the characteristics of the features (= data we use for training), preparation of training set, and choice of Kernel functions done for this study.

2.1. Feature selection

We need data to get the classifier we want through the SVM. In machine learning, data is typically called features and these features are used to find a hyperplane that can distinguish the classes. Therefore, feature selection is an important step for making a classifier using SVM. Features are chosen from known plasma diagnostics which indicate the occurrence of L-H transition, such as $\boldsymbol{D_\alpha}$ signal drop, increase of line integrated electron density $\overline{\boldsymbol{n_e}}$, as well as electron and ion temperature [8]. In addition, the features should satisfy the requirement that they are available for real-time use. That is because the goal of our research is to help improve real-time density control efficiency in plasma operation.

In this work, the SVM classifier is trained with as few as possible features because its final goal is real-time calculation which requires

* Corresponding author at: National Fusion Research Institute, 113 Gwahak-ro, Yuseong-gu, 34133, Republic of Korea and Korea University of Science and Technology, 217 Gajaeong-ro, Yuseong-gu, 34113, Republic of Korea.
E-mail address: hahn76@nfri.re.kr (S.H. Hahn).

https://doi.org/10.1016/j.fusengdes.2017.12.011
Received 9 June 2017; Received in revised form 14 November 2017; Accepted 11 December 2017

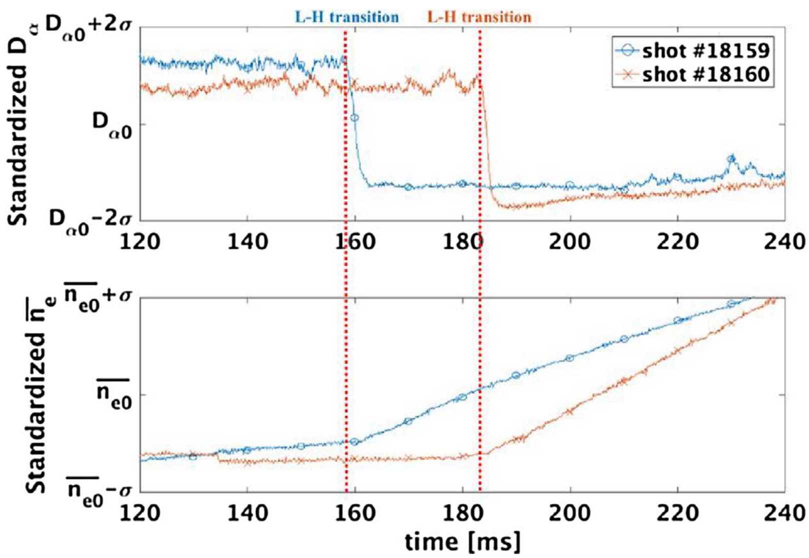

Fig. 1. (top) Amplitude change and drop timing of D_α signal in typical KSTAR discharges and (bottom) increase of standardized line-averaged density at the moment of the L-H transition in KSTAR plasmas. The $D_{\alpha 0}$ and $\overline{n_{e0}}$ are the mean values of each measurement and the σ is the standard deviation of the D_α and $\overline{n_e}$.

fast computation speed. As the number of features increases, the number of support vectors, which is the closest vectors to the classifier [4], increases hence the amount of calculation increases. In the SVM classifier, the number of support vectors affects calculation speed because the whole matrix of the support vectors should be computed for single test data. Therefore, the proper number of features should be selected to guarantee the accuracy of predictions above a certain level.

The $\boldsymbol{D_\alpha}$ signal, shown in the top of Fig. 1, is the most important feature since its characteristic intensity drop pattern indicates the edge temperature increases due to improved confinement as edge turbulence is suppressed. [8]. Since the height of the signal drop is not constant, and varies shot by shot, this $\boldsymbol{D_\alpha}$ signal feature cannot be classified by a simple linear classifier. The characteristic time scale of the $\boldsymbol{D_\alpha}$ drop at the L-H transition is usually about 2–4 milliseconds in KSTAR, and usually it is accompanied by repeating burst events called as Edge Localized Mode(ELM) [9] within several hundred milliseconds. In order to detect the L-H transition and to properly start operating an algorithm for density control, it is desired that the computational time is faster than the time scale of dropping the $\boldsymbol{D_\alpha}$ signal. This is another reason why we consider the SVM for the classifier: Considering the time until the first Edge Localized Mode (ELM) burst occurs after the L-H transition, the classifier should give the status (L or H) within 150–300 milliseconds so that the corresponding controller could respond to the status change.

Another feature under consideration is the line-averaged electron density measurement $\overline{n_e}$ by millimeter wave interferometer(MMWI) [10]. As shown in the bottom of Fig. 1, the increase of the electron density is big enough to distinguish the H-mode from the L-mode due to the edge pedestal formation and improved particle confinement. Other features, such as pressure or electron temperature, did not have enough distinctive temporal profile that can train the SVM.

As shown in Fig. 1, there is a change that can distinguish H-mode from L-mode after the L-H transition occurs in both features. The D_α drop occurs in a relatively short time interval whereas the density signal increases more slowly, although the beginning of the changes occurs almost at the same time within the available time resolution (10 kSamples/s).

2.2. Preparation of data set

Once the features are selected, the preparation of training/testing data set is done as the next step. Our area of interest is ± 100 ms around the moment when the L-H transition is seen. We extracted 2001 samples/shot from the 10KHz-sampled D_α measurement. Since the sampling rate of the line integrated electron density $\overline{n_e}$ is originally 100KHz, the density measurement has been downsampled to the same sampling (10KHz) as the D_α signal.

Since the quality of the diagnostics can vary shot to shot, we prepared separate data sets for a single feature model and a two-features model. The single feature model uses only the D_α signal. 200 independent shots are prepared for this model, and 70% of the data set are used for training and 30% of the data set are used for testing the trained model. In other words, 140 shots are used as training set and 60 shots as test set.

In the model with two features using both the D_α signal and line-averaged electron density $\overline{n_e}$, the total number of prepared shots is 197. Since the number of samples obtained from one shot is 2001, 394,197 samples were provided from those 197 shots. In these samples, 278,319 samples from 139 shots were used as training set and 116,058 samples from the remaining 58 shots were used as test set.

2.3. Non-linear classifier, Kernel function, and formula for the real-time detection

The classification of the data should be done by a non-linear classifier. As shown in Fig. 2, the two-scattered data cannot be separated by a linear classifier.

The form of a classifier is as the following:

$$\boldsymbol{d}(\boldsymbol{x}) = \sum_{x_k \in Y} \boldsymbol{\alpha_k y_k K}(\boldsymbol{x_k}, \boldsymbol{x}) + \boldsymbol{b}$$

where $\boldsymbol{x_k}$ support vectors, $\boldsymbol{\alpha_k}$ Lagrange multiplier, $\boldsymbol{y_k}$ the label of support vectors, $\boldsymbol{b}$ bias and $\boldsymbol{K}(\boldsymbol{x_k}, \boldsymbol{x})$ the Kernel function that has to be figured out for the accuracy of separation.

In order to find a suitable kernel function for our non-linear SVM classifier, we evaluated our model through 5-fold cross-validation, which is one of the cross-validation techniques in machine learning. In the 5-fold cross-validation, the original training set is randomly divided into 5 subsets. Of the 5 subsets, one subset is used for test data and the remaining 4 subsets are used for training data. The validation process is repeated 5 times, changing the subsets. The error of the cross-validation is obtained by averaging the results of 5 times. The advantages of this technique are that the correlation between the original training data is less affected since the training data is randomly partitioned and the calculation speed of cross-validation is fast.

Through the 5-fold cross-validation, we have validated the kernel function K with quadratic, cubic, and radial basis function (RBF) in the form