

B. Cannas, R.S. Delogu, A. Fanni, P. Sonato, M.K. Zedda  
and JET EFDA contributors

# Support Vector Machines for Disruption Prediction and Novelty Detection at JET

“This document is intended for publication in the open literature. It is made available on the understanding that it may not be further circulated and extracts or references may not be published prior to publication of the original when applicable, or without the consent of the Publications Officer, EFDA, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK.”

“Enquiries about Copyright and reproduction should be addressed to the Publications Officer, EFDA, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK.”

# Support Vector Machines for Disruption Prediction and Novelty Detection at JET

B. Cannas<sup>1</sup>, R. S. Delogu<sup>1</sup>, A. Fanni<sup>1</sup>, P. Sonato<sup>2</sup>, M. K. Zedda<sup>1</sup>  
and JET-EFDA contributors\*

<sup>1</sup> *Electrical and Electronic Engineering Dept.-University of Cagliari, Piazza D'Armi, 09123, Cagliari, Italy.*

<sup>2</sup> *Consorzio RFX, Associazione Euratom-ENEA sulla Fusione, Corso Stati Uniti, 4, I-35127 Padova, Italy*

*\* See annex of J. Pamela et al, "Overview of JET Results",*

*(Proc. 20<sup>th</sup> IAEA Fusion Energy Conference, Vilamoura, Portugal (2004)).*

Preprint of Paper to be submitted for publication in Proceedings of the  
SOFT Conference,  
(Warsaw, Poland 11th – 15th September 2006)



## **ABSTRACT**

In the last years there has been a growing interest on neural approaches to disruption prediction. The drawback of these approaches is that the system performance could deteriorate once it is on-line. This could be the case for a disruption predictor for JET, where new plasma configurations might present features completely different from those observed in the experiments used during the training phase. This 'novelty' can lead to incorrect behaviour of the network. A Novelty Detection method, which determines the novelty of the input of the prediction system, can be used to assess the network reliability.

This paper presents a Support Vector Machines disruption predictor for JET, wherein multiple plasma diagnostic signals are combined to provide a composite impending disruption warning indicator. In a Support Vector Machine the analysis of the decision function value gives useful information about the novelty of an input and, on the reliability of the predictor output, during on-line applications. Results show the suitability of Support Vector Machines both for prediction and novelty detection tasks at JET.

## **1. INTRODUCTION**

Disruptions pose serious problems to the integrity and the lifetime of a tokamak. For this reason in the last 15 years, there have been several studies for disruptions prediction, most of which use neural networks [1-3].

One of the major drawbacks of this approach is that the network performance normally deteriorates when new plasma configurations are presented to the network. In fact, a network, which is trained to discriminate between inputs coming from a set of distributions, could produce unreliable output when input data comes from an entirely new distribution. Improvements might be possible using Novelty Detection (ND) techniques [4]. In the on-line application, the Novelty Detection should be used to assess the reliability of the network output, i.e., samples having a low confidence have to be discarded and used off-line to update the disruption predictor.

The Novelty Detection methods proposed in the literature are based on both statistical and neural network approaches [5, 6].

In [7], a Novelty Detection system was implemented to integrate the neural disruption predictor proposed for JET. One of the limits of this approach is that the two systems work separately, and the novelty detector is able to capture only the topological information present in the database, and not information which the neural network learnt about that database. In this paper, both the prediction and the novelty detection tasks are performed by the same system using a Support Vector Machine (SVM).

## **2. SUPPORT VECTOR MACHINES**

An SVM is a supervised algorithm proposed first by Vapnik [8]. The algorithm addresses the general problem of discriminating between two classes of n-dimensional vectors.

The SVM [9] is built by detecting an optimal separating hyper-plane, which maximizes the margin between itself and the closest training data. An optimal separating hyper-plane separates the

data in two classes. In case of non linear decision functions, the SVM projects the input vector  $x$  into a high-dimensional feature space  $H$  and constructs the optimal separating hyper-plane in this space. The optimal hyper-plane is obtained by solving the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & E(w) = \frac{1}{2} w w^T + C \sum_{i=1}^l \xi_i \\ \text{subject to} \quad & \begin{cases} y^i (w^T \Phi(x^i) + b) \leq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \quad i = 1, \dots, l \end{aligned}$$

where  $w$  and  $b$  are the coefficients of the optimal hyper-plane,  $C$  is a error penalty, and  $\xi_i$  are parameters handling non separable inputs. The index  $i$  labels the  $l$  training cases;  $x^i$  is the  $i^{th}$  training input pattern, and  $y^i \in \{-1, +1\}$  is the corresponding class label. The kernel  $\Phi$  is used to transform data from the input to the higher dimensional feature space.

There are different kernels that can be used in SVM models, e.g.:

- polynomial  $\Phi(x^i, x^j) = (\gamma x^i \cdot x^j + \vartheta)^p$ ,
- gaussian radial basis  $\Phi(x^i, x^j) = e^{-\|x^i - x^j\|^2 / 2\sigma^2}$

In particular, if the kernel function is a radial basis function, the corresponding feature space is a Hilbert space of infinite dimension. The images of the input are always linearly separable in the feature space [10].

It should be noted that all the parameters should be carefully chosen to avoid over fitting. Once an SVM is trained, the class membership of a given test pattern  $x$  is determined depending on the sign of the decision function:

$$f(x) = w^T \cdot \Phi(x) + b.$$

The feature making SVMs very attractive is that classes non linearly separable in the original space can be linearly separated in the higher dimensional feature space. Thus, SVM is capable to solve complex non linear classification problems.

### 3. NOVELTY DETECTION TECHNIQUE

Novelty Detection consists of identifying new or unknown data that a machine learning system is not aware of during the training phase. Thus, novelty detection is one of the fundamental requirements of a good classification or prediction system. In fact, actual data may contain patterns belonging to operational regions not explored when the learning system was developed. This could be the case of the disruption predictor presented in this paper, where new plasma configurations might present features completely different from those observed in the experiments selected for the training set. This 'novelty' can lead to incorrect behaviour of the SVM predictor.

In the last ten years novelty detection has acquired increasing attention, and a number of techniques have been proposed and investigated to address it. In [5] and [6] the authors highlighted that it is not possible to identify a priori a single best model, and the success of a novelty detection technique mainly depends on the statistical properties of data handled.

Both statistic and neural clustering methods can be used for novelty detection tasks.

In this paper, the novelty detection is performed by inspecting the SVM decision function value. As values of the decision function close to +1 or to -1 correspond to a correct answer of the SVM predictor, two bands  $[1 - \Delta_a, 1 + \Delta_b]$  and  $[-1 + \Delta_a, -1 - \Delta_b]$  can be introduced to identify novel samples. In particular, during the test phase, a sample will be labelled as not novel if its decision function has a value belonging to that bands, and conversely.

The values  $\Delta_a$  and  $\Delta_b$  are chosen such that the decision function falls within that bands for the 90% of the training samples.

In Fig. 1, the decision function values are reported for a test pulse. In the same figure the two bands are highlighted. Samples falling out of that bands will be labelled as novel by the novelty detector, and the output of the predictor will be inhibited.

#### 4. DATABASE SELECTION

The diagnostic signals for training and testing the SVMs were selected in the interval of JET Pulses No 47830 - 57346. The database consists of 172 disruptive pulses and 102 non disruptive pulses.

The pulses belong to different classes of disruption: Mode Lock, Density Limit, high Radiated Power, H-mode/L-mode transition, Internal Transport barrier [11] [12].

The discharges included in the network database satisfy the following requirements:

- Plasma current  $I_{PLA} > 1.5$  MA;
- X-point configuration;
- Stationary (Flat-top) plasma current profile.

Discharges with IPLA below 1.5MA were discarded as they generally have little impact on subsequent conditioning and operation of the device.

Nine diagnostic signals have been selected to describe the plasma regime during the discharge flat-top. These signals represent the input of the SVM. The choice of the signals takes into account physical considerations and the availability of real-time data. Moreover, previous experiences on disruption prediction confirm the appropriateness of the chosen input variables [2].

The sampling time is 20ms in order to allow the synchronization among different acquisition systems [12].

Table I shows the selected diagnostic signals.

In order to train, validate and test the network performance [10], the pulses in the database have been divided in three sets: the training set consists of 69 disrupted pulses, the validation set consists of 17 disrupted pulses, while the test set consists of 86 disrupted pulses and 102 successful pulses.

## 5. THE SVM PREDICTOR/NOVELTY DETECTOR

The predictive system structure consists of two blocks mutually connected: a Self Organising Map (SOM) and an SVM (see Fig.2).

During the training phase, a SOM performs a clustering procedure and an SVM is trained to give the alarm in case of impending disruption. When the training procedure is completed, the previously introduced bands, which discriminate between novel and not novel samples, are calculated.

In order to build the training set for the SVM predictor, 86 SOMs have been constructed, one for each pulse in the training and validation sets. Each SOM is used to identify the precursor phase of the corresponding disrupted pulse, i.e., to discriminate between ‘safe’ samples and ‘disrupted’ samples containing information about the disruption proximity. Moreover, the SOM is used for data reduction, i.e., only one safe sample for each cluster is selected to build the SVM [13].

The SVM is trained to classify safe and disrupted samples respectively labelled as +1 and -1. During the on-line application, the SVM is fed with all the samples of a pulse, and, for each of them, it returns a label equal to -1 or +1.

In this paper the performance of the prediction system is evaluated in terms of Percentage of False Alarms (PFA) and Percentage of Missed Alarms (PMA), where PFA is defined as the ratio between the number of non disruptive pulses predicted by the system as disruptive shots, and the total number of non disruptive pulses, in per cent, while PMA is defined as the ratio between the number of disruptive pulses predicted as non disruptive pulses, and the number of disruptive pulses, in per cent. Note that a disruption prediction is considered successful if the system is able to correctly predict the disruption up to 100ms prior to the disruption time.

Furthermore, the value of the decision function is used to assess the reliability of the SVM output: SVM output related to a test sample, which presents a decision function value falling out of the two bands  $[1 - \Delta_a, 1 + \Delta_b]$  and  $[-1 + \Delta_a, -1 - \Delta_b]$ , has to be rejected, that sample is recognized to be novel, and it can be used to update the disruption predictor.

## 6. RESULTS

The SVM has been trained using the OSU SVM Classifier Matlab Toolbox (ver. 3.00) based on the version 2.33 of LIBSVM [14]; LIBSVM is a software library for classification and regression by means of Support Vector Machines that implements the training algorithms developed by Vapnik.

Several training sessions have been performed in order to select the SVM parameters. The best results have been obtained with the Radial Basis Function Kernel with  $\sigma = 18.7$  and error penalty  $C = 40$ .

Table II shows the performance of the SVM predictor for the training, the validation, and the test set respectively, in terms of PFA and PMA.

Note that, for each pulse of the training and validation sets, the errors have been calculated considering the whole sequences of the samples rather the samples selected by the SOM during the training phase.



The network performance is excellent in terms of false alarms, but the percentage of missed alarms is quite high.

It has to be pointed out that for 9 disruptive pulses the system triggers the alarm too early (1s) with respect to the disruption time.

Table III reports the results obtained considering a successful prediction if the alarm is triggered up to 80, 60, 40ms before the disruption time  $t_D$ . Since PFA is 0% up to 100ms, only the PMA are reported.

The performance of the SVM predictor, integrated with the ND block, is reported in Table IV in terms of PMA and PFA calculated on a reduced test set obtained by the previous test set by discarding the pulses labelled as novel. A sample is considered novel if  $1 - \Delta_a < |f(\mathbf{x})| < 1 + \Delta_b$  where  $\Delta_a = 0.5$  and  $\Delta_b = 1$ .

The Novelty Detector influences the predictor behaviour only in case of disruption alarm. In particular, if the network triggers the alarm for a sample considered 'novel', the alarm will be rejected. In particular 14 of the 23 missed alarms triggered by the SVM predictor are labelled as novel by the ND. Moreover, 5 of the 9 disruptive pulses, which have not been considered as correctly predicted (the system triggered the alarm too much in advance with respect to the disruption time) are labelled as novel too. Note that, even if the alarm is rejected by the ND, for a safe operation of the machine, the discharge should be managed by the human interventions, stopping the discharge, especially under conditions where disruptions can not be tolerated.

It is worth noting that some disruptive pulses, correctly predicted by the SVM, have been labelled as novel by the ND. Hence, although the number of FAs and MAs decreases, the discrimination capability of the system in the on-line application reduces, but the robustness and the reliability of the system increase.

## CONCLUSION

The unavoidable ageing of a neural prediction system is important for the machines, such as JET, where new the plasma configurations are explored. So, it is crucial to have a system able to measure the reliability of the network output and to automatically update the network to include plasma configurations not used during the training phase.

An SVM is proposed, which integrates prediction and novelty detection capabilities in a unique system.

The SVM predictor shows a null percentage of false alarms, while the percentage of missed alarms is quite high. Using the knowledge acquired during the training phase of the predictor, the system is able to detect the novelty of new pulses increasing the performance of the entire system.

In particular, the novelty detector is able to justify many of the missed alarms of the predictor as they are recognized as belonging to new regions of the operational space.

## REFERENCES

- [1]. G. Pautasso, C.Tichmann, et al. “On-line prediction and mitigation of disruption in ASDEX Upgrade”, Nucl. Fusion 37, pp. 100-108, 2002.
- [2]. B. Cannas, A. Fanni, G. Sias, P. Sonato, M.K. Zedda, and JET EFDA contributors “Neural approaches to disruption prediction at JET”, 31th EPS Conf. on Plasma Phys., 28G, p-1.167, London, UK, 2004.
- [3]. R. Yoshino “Neural-net disruption predictor in JT-60U”, Nucl. Fusion 43, pp. 1771-1786, 2003.
- [4]. C. M. Bishop, “Novelty detection and neural network validation”, Proc. of the IEE Conf. on Vision and Image Processing, pp 217-222, 1994.
- [5] M. Markou, S. Singh, “Novelty Detection: a review – part 1: statistical approaches”, Signal Processing, 83, pp 2481-2497, 2003.
- [6]. M. Markou, S. Singh, “Novelty Detection: a review – part 2: neural network based approaches”, Signal Processing, 83, pp. 2499-2521, 2003.
- [7]. B. Cannas, A. Fanni, P. Sonato, M.K. Zedda, and JET EFDA contributors to the EFDA-JET workprogramme “Novelty Detection for on-line disruption prediction systems”, 32th EPS Conference on Controlled Fusion and Plasma Physics, Tarragona, Spain, 27 June - 1 July 2005.
- [8]. Vapnik V. “Statistical Learning Theory”, Wiley-Interscience, New York 1998.
- [9]. Burges C.J.C. “A tutorial for Support Vector Machines for Pattern Recognition”, Data Mining and Knowledge Discovery, 2, pp 121-167, 1998.
- [9] Cristianini N and Shawe-Taylor “An introduction to Support Vector Machines”, Cambridge University Press, 2000.
- [10] M.K. Zedda, T. Bolzonella, B. Cannas, A. Fanni, D. Howell, M.F. Johnson, P. Sonato and JET EfdA contributors, “Disruption classification at JET with neural techniques”, 30th EPS Conference on Contr. Fusion and Plasma Phys., Vol. 27A, p-2.93. St. Petersburg, 7-11 July 2003.
- [11]. B. Cannas, F.Cau, A. Fanni, P. Sonato, M.K. Zedda and JET-EFDA contributors “Automatic disruption classification at JET: Comparison of different pattern recognition techniques” Nucl. Fusion 46, 699-708, 2006.
- [12]. B. Cannas, A. Fanni, G. Sias, P. Sonato, M.K. Zedda, and JET EFDA contributors “Neural approaches to disruption prediction at JET”, 31th EPS Conference on Controlled Fusion and Plasma Physics, Vol.28G, p-1.167, London, UK, 28 June – 2 July 2004.
- [13]. Chang C.C, Lin C.J.. 2001 LIBSVM a library for support vector machines Software available at:<http://www.csie.ntu.edu.tw>.

:

<b>Signal name</b>	<b>Units</b>
1. Plasma current	[A]
2. Locked Mode	[T]
3. Radiated power	[W]
4. Plasma Density	[1/m <sup>3</sup> ]
5. Input Power	[W]
6. Internal Inductance	
7. Safety factor	
8. Poloidal Beta	
9. Plasma centroid vertical position	[m]

Table I: Diagnostic signals.

	<b>Training Set</b>	<b>Validation Set</b>	<b>Test Set</b>
PMA	0%	23.5%	26.7%
PFA	–	–	0%

Table II: SVM Predictor performance in terms of Percentage of False Alarms (PFA), and Percentage of Missed Alarms (PMA).

	<b>t<sub>D</sub>-80ms</b>	<b>t<sub>D</sub>-60ms</b>	<b>t<sub>D</sub>-40ms</b>
PMA	26.7%	25.5%	17.4%

Table III: SVM Predictor performance in terms of Percentage of Missed Alarms (PMA) up to different time instants.

	<b>Test Set</b>
PMA	13.8%
PFA	0%

Table IV: SVM Predictor/Novelty Detector performance in terms of Percentage of False Alarms (PFA), and Percentage of Missed Alarms (PMA).

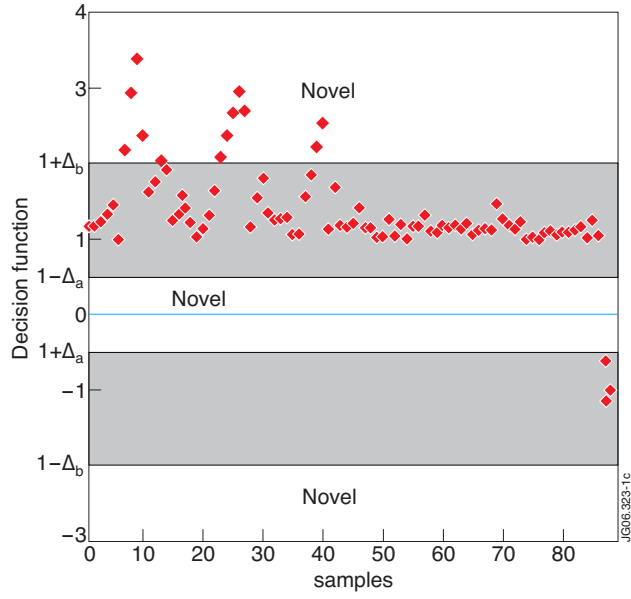


Figure 1: Decision function for a test pulse. The grey regions represent the bands discriminating novel and not novel samples.

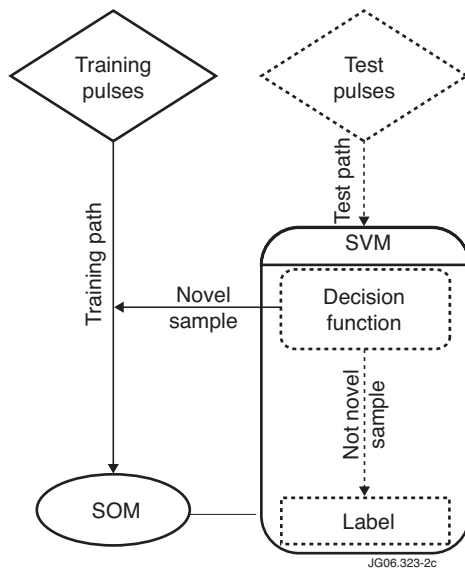


Figure 2: Architecture of the Predictor/ Novelty Detector system. Continuous lines: training path; dashed line: test path.