 Printed in China and the UK Plasma Science and Technology

Plasma Sci. Technol. 24 (2022) 124003 (12pp) https://doi.org/10.1088/2058-6272/ac9e46

Overview of machine learning applications in fusion plasma experiments on J-TEXT tokamak

Wei ZHENG (郑玮)[1], Fengming XUE (薛凤鸣)[1,2], Chengshuo SHEN (沈呈硕)[1], Yu ZHONG (钟昱)[1], Xinkun AI (艾鑫坤)[1], Zhongyong CHEN (陈忠勇)[1], Yonghua DING (丁永华)[1], Ming ZHANG (张明)[1], Zhoujun YANG (杨州军)[1], Nengchao WANG (王能超)[1], Zhichao ZHANG (张智超)[1], Jiaolong DONG (董蛟龙)[1], Chouyao TANG (唐畴尧)[1] and Yuan PAN (潘垣)[1]

[1] International Joint Research Laboratory of Magnetic Confinement Fusion and Plasma Physics, State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China
[2] Institute of Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China

E-mail: zhengwei@hust.edu.cn

Received 31 August 2022, revised 30 September 2022
Accepted for publication 27 October 2022
Published 30 November 2022

Abstract

Machine learning research and applications in fusion plasma experiments are one of the main subjects on J-TEXT. Since 2013, various kinds of traditional machine learning, as well as deep learning methods have been applied to fusion plasma experiments. Further applications in the real-time experimental environment have proved the feasibility and effectiveness of the methods. For disruption prediction, we started by predicting disruptions of limited classes with a short warning time that could not meet the requirements of the mitigation system. After years of study, nowadays disruption prediction methods on J-TEXT are able to predict all kinds of disruptions with a high success rate and long enough warning time. Furthermore, cross-device disruption prediction methods have obtained promising results. Interpretable analysis of the models are studied. For diagnostics data processing, efforts have been made to reduce manual work in processing and to increase the robustness of the diagnostic system. Models based on both traditional machine learning and deep learning have been applied to real-time experimental environments. The models have been cooperating with the plasma control system and other systems, to make joint decisions to further support the experiments.

Keywords: machine learning, disruption prediction, diagnostics data processing, J-TEXT

(Some figures may appear in colour only in the online journal)

1. Introduction

Machine learning has been the subject of extensive studies in scientific research lately. Recent advances in computing power have promoted machine learning, especially deep learning applications in many fields. Today's deep learning models are powering many commercial applications with higher accuracy and speed beyond the human level. A large amount of scientific research is data-heavy, which means a huge amount of data is produced through experiments and simulations. Thus, machine learning is a quite suitable tool in scientific research. It can not only free researchers from tedious data processing tasks but it can also discover new phenomena or even new laws. Deep learning methods and applications have already been widely used in biology [1], materials [2], high energy physics research [3], as well as other fields.

Fusion plasma research is another field where machine learning has bloomed in recent years. Machine learning was applied to fusion plasma research a long time ago. In 1996, researchers from TEXT-U tried to use a preliminary neural network to predict disruptions [4]. Limited by the computing power and data volume available at the time, the neural network which was applied in disruption prediction used only a few layers and was only able to predict some classes of disruption. However, it revealed the ability of neural networks to predict disruptions for the first time. With the development of computing power, researchers from Harvard and the Princeton Plasma Physics Laboratory (PPPL) from Princeton University, have used deep neural networks and have achieved promising results in disruption prediction with high accuracy [5]. Today, disruption prediction remains the most active subject in machine learning fusion plasma research. Another emerging application of machine learning fusion plasma research is reinforcement learning in plasma control. The DeepMind team together with the TCV tokamak have developed a reinforcement learning plasma control architecture [6]. The plasma controller is end-to-end, which means that no more equilibrium design is needed when designing a new plasma configuration. This will make it much easier for tokamak operators. In addition, there are also many other machine learning applications in tokamak diagnostics data processing and plasma simulation [7–11].

Machine learning in fusion plasma research is a major subject on J-TEXT tokamak. The article will give a comprehensive review of machine learning research on J-TEXT. One of the most investigated subjects is machine learning-based disruption prediction, which will be presented in section 2. In section 3, machine learning applications on diagnostics data processing are presented. In section 4, machine learning methods, which were implemented in real-time and applied in actual J-TEXT experiments, are presented. A summary is described in the last section.

2. Machine learning in disruption prediction

Disruption is a catastrophic loss of tokamak plasma confinement and is a great obstacle to overcome for successful and stable operations in future tokamaks [12–17]. Disruptions will cause strong electromagnetic forces, heat loads, and runaway electrons during the thermal and current quench phase, which will do great harm to the structure of the tokamak. Thus, disruptions at high-performance discharge are unacceptable, especially for future fusion reactors. In order to alleviate the aforementioned harm caused by disruptions, disruption prediction, prevention and mitigation are of great importance. Basically, all current disruption prevention and mitigation systems require a time scale of several milliseconds to be fully effective. The key for the systems to take effect is to accurately predict the impending disruptions with sufficient warning time for the actuators to work. However, the physical mechanism of disruption is still not clear enough. Rule-based methods are not likely to reach a promising result and are not fully capable of using all information contained by the diagnostics of the tokamak. Data-driven methods seem to be able to handle the problem. Based on a large amount of data and proper machine learning techniques, the methods are able to recognize various kinds of disruptions at high accuracy and are even able to discover relations to disruptions that were not previously found.

Research on predicting disruption has been carried out since the 1990s. Except for the neural network disruption predictor realized on the TEXT-U tokamak in 1996 [4], JET also explored disruption prediction methods based on neural network and novelty detection techniques [18]. A neural network disruption predictor designed for the high-beta limit was composed on the DIII-D tokamak in 1997 [19]. The JT-60U tokamak realized a neural network-based disruption predictor in 2003 [20]. A new criterion for predicting disruptions was put forward in HL-2A in 2006 [21]. Machine learning-based disruption prediction is also one of the main subjects on J-TEXT. With years of development and further research, disruption prediction has become a featured advantage on J-TEXT.

2.1. Predicting disruptions caused by locked modes

The first attempt to predict disruptions was carried out in 2013 on J-TEXT. A neural network model was established to predict disruptions caused by locked modes [22]. The model took Mirnov coils and locked mode signals as inputs to better focus on locked mode-induced disruptions. The regression model was based on a fully connected neural network and took three hidden layers' topology, which is illustrated in figure 1. According to the statistics, a success rate of 90% was reached and only two discharges disrupted induced by locked modes were not predicted due to a short flattop phase. For intrinsically locked mode disruptions, a warning time of about 1 ms ahead of the locking time could be given. For disruptions caused by resonant magnetic perturbations (RMPs) locked modes [23], a warning time of 10 ms could be given. This work for the first time validated the feasibility of predicting disruptions on the J-TEXT tokamak and proved that neural network-based methods are capable of detecting disruption precursors.

2.2. Predicting disruptions caused by density limit

Disruption induced by density limit is another major type of disruption on the J-TEXT tokamak. A neural network-based density limit predictor was first developed for predicting disruptions induced by density limits [24]. By targeting a specific disruption type, it was easier for the neural network model to detect the related disruption precursors. In this work, rather than predicting the probability of an incoming disruption, the neural network model tried to estimate the current density limit value so that the plasma control system (PCS) could use the estimated density limit as a reference to avoid the density limit while operating as close as possible to the

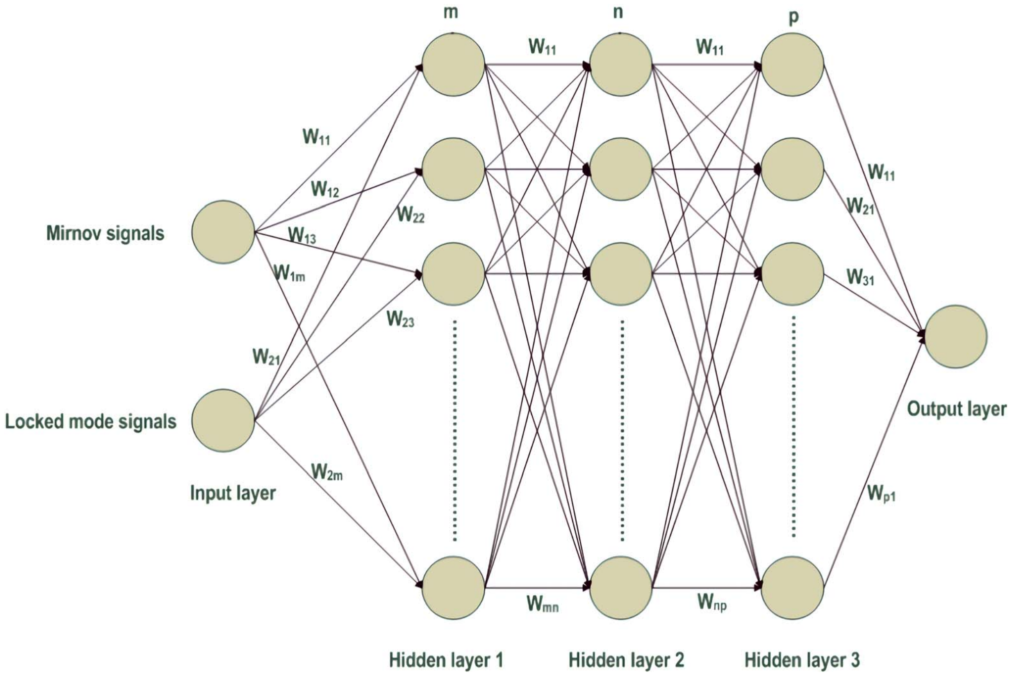

Figure 1. Topology of the neural network model designed for predicting locked modes induced disruptions. The model used Mirnov signals and other locked mode signals as input, and alarming of the locked mode as output. (Reproduced from [22].

limit. The method performed well in the offline test, successfully predicting 82.8% of the density limit disruptions with a warning time of over 4 ms. Moreover, if the predicted density limits were fed to the density feedback controller to optimize the experimental density, more than 95% of density limits induced disruptions could be avoided. The results obtained showed that the designed network was suitable for online density limit avoidance and could be implemented in the J-TEXT tokamak control system to help avoid or mitigate density limit disruptions with improvement in the warning time prior to disruptions.

Further investigation on predicting disruptions caused by density limits was carried out [25]. The method labeled all non-density limit disruptions as negative samples, expanding the training set dramatically. It also no longer used a simple fully connected neural network but designed and applied a two-stage hybrid neural network. The first stage was a time series neural network used to predict future density, while the second was a fully connected neural network used to predict the density limit disruption probability. The two networks were trained separately and were combined as a whole at the inference stage. The predicted density which was given by the first neural network was fed into the second neural network as one of the inputs. With the improved two-stage hybrid neural network, the warning time performance was significantly improved by about 30 ms ahead, and the true positive rate was increased by 5%–10% at the same time. The performance comparisons between the hybrid neural network and the neural network without predicting the density limit are shown in figure 2. Furthermore, to reach a higher true positive rate and a lower false positive rate, the inputs of the second stage neural network were reselected to cover profile information of the radiation signals. A performance comparison between inputs with and without profile information is shown in figure 3. The true positive rate was greatly improved using the information from the radiation profile. The neural network finally reached a true positive rate of over 90% and a false alarm rate of less than 10%, with an average warning time longer than 30 ms. The model was then realized in the real-time environment in J-TEXT experiments to avoid density limit disruptions, which would be described in detail in section 4. The work further proved that neural network methods were capable of extracting and predicting disruption precursors, and first implemented disruption prediction in the real-time experimental environment to help make joint decisions with the plasma control system (PCS) to realize disruption avoidance and mitigation.

2.3. Predicting disruptions using a deep learning-based anomaly detection technique

As machine learning-based disruption prediction methods need abundant data for training and inferencing, it is necessary to develop a database built with disruption-related information. The database should also be able to generate clean and labeled datasets conveniently. With this aim, a database designed specifically for machine learning-based disruption predictors was developed [26]. The structure of the database is shown in figure 4. All data from any tokamak devices could be transferred to a standard format and saved as HDF5 files. Labels as well as other metadata of each discharge could be stored in MongoDB. Information about each discharge was able to be added, removed, checked, and changed with ease.

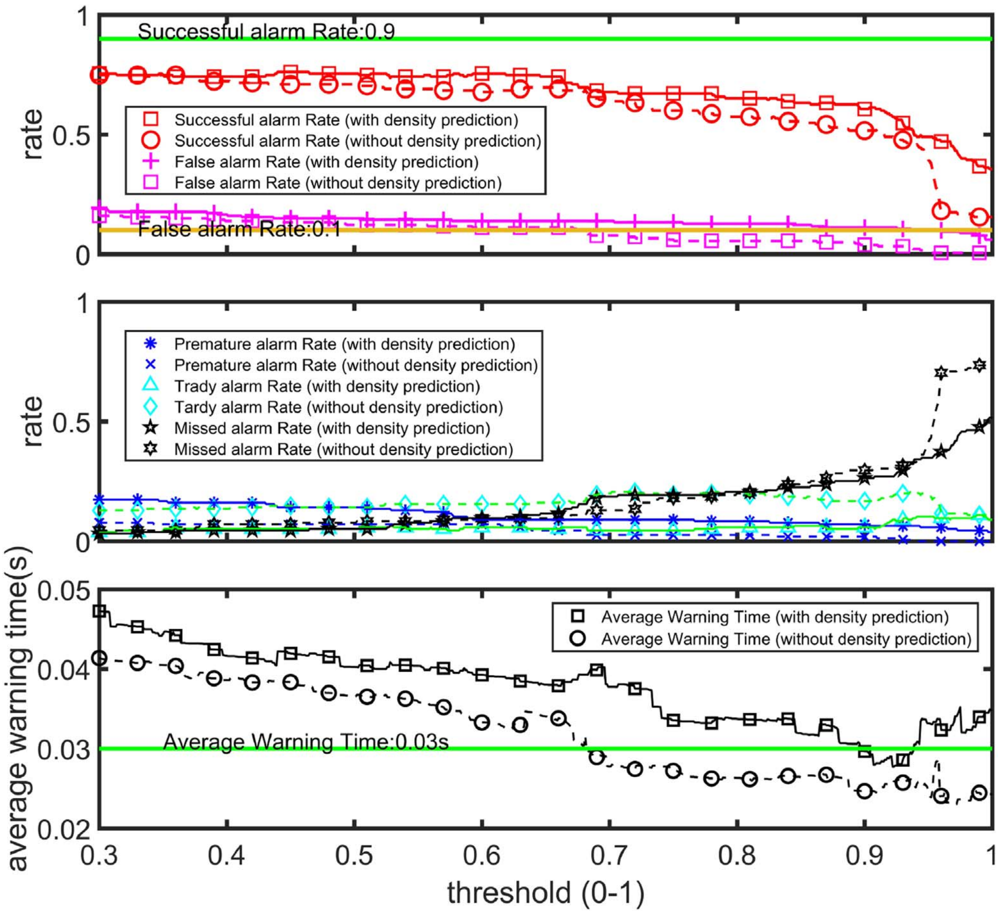

Figure 2. Overall comparisons of the two neural networks with and without density prediction. In the upper part, the red and magenta lines represent successful and false alarm rates with and without density prediction, respectively. In the middle part, the blue, cyan, and black lines represent premature, tardy, and missed alarm rates with and without density prediction, respectively. In the lower part, the black lines indicate average warning times with and without density prediction. (Reproduced from [25].

Based on the disruption database above, disruption prediction based on deep learning anomaly detection was investigated [27]. One problem with disruption prediction is that, for future tokamaks, it is almost impossible to obtain enough disruptive discharges for training a data-driven model, because large tokamaks can hardly bear any disruptive discharges without damaging themselves. Thus, the method applied one-class data along with its label for training, which was the non-disruptive class. Considering that most discharges with locked mode will finally lead to disruptions, a hybrid neural network was designed to predict the value of the signals which were related to locked modes 20 ms later from current inputs. The model combined CNNs (Convolutional Neural Networks) with LSTMs (Long-Short Term Memory neural networks) together. The 2D convolution layers were designed to extract high-frequency features from the Mirnov array, while the features were then concatenated with other low-frequency signals and were fed into LSTM to predict future information of the locked mode. With the hypothesis that most disruptive discharges are induced by locked mode, which hardly appears in non-disruptive discharges, the differences in the predicted locked mode between disruptive and non-disruptive discharges were supposed to be significant. An outlier factor was calculated based on this principle to identify whether the discharge is disruptive or not. As a result, the method reached a true positive rate of 83% and a false positive rate of 18%, with an average warning time of 36 ms. Typical predictions of disruptive and non-disruptive discharges are shown in figure 5. Different from the previously mentioned models, which predicted a specific type of disruption, the method was designed to predict general types of disruptions on J-TEXT. Moreover, it used only the non-disruptive discharges for training, which attempted to solve the problem of not having enough disruptive discharges for training in future tokamaks.

2.4. Predicting disruptions across different tokamaks using a transfer learning technique

Research on predicting disruptions across different tokamaks was carried out and promising results were obtained on the J-TEXT and the EAST tokamak [28]. Basically, a disruption predictor consists of a feature extractor and a classifier. The key is to design a feature extractor that is able to extract general features across different tokamaks. A deep learning-based

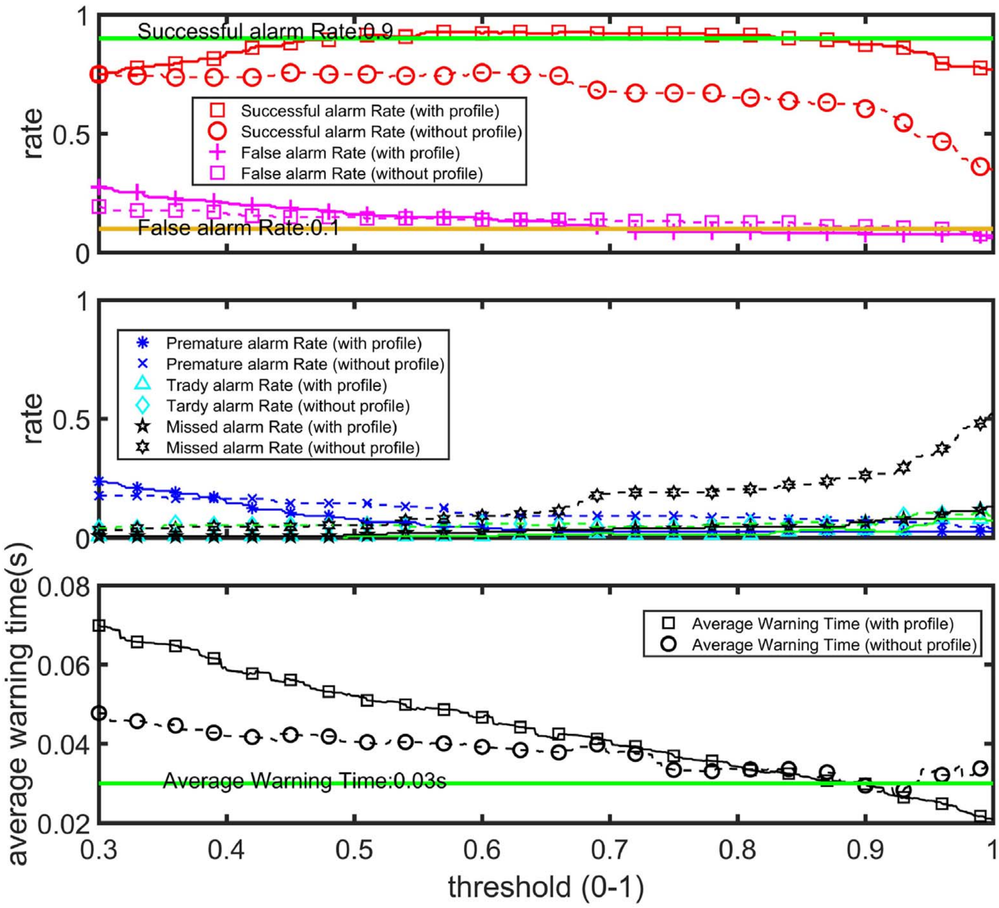

Figure 3. Overall comparison of the two neural networks with and without profile information. In the upper part, the red and magenta lines represent successful and false alarm rates with and without density prediction, respectively. In the middle part, the blue, cyan, and black lines represent premature, tardy, and missed alarm rates with and without density prediction, respectively. In the lower part, the black lines indicate average warning times with and without density prediction. (Reproduced from [25]. © IOP Publishing Ltd. All rights reserved).

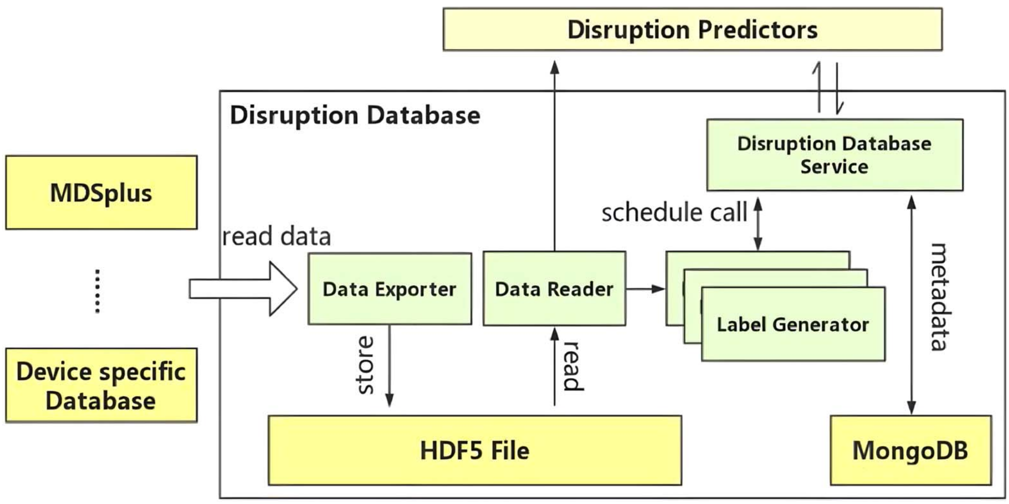

Figure 4. The structure of the disruption database. Device specific database will be converted to a standard format and stored as an HDF5 file. Different plugins will be loaded in the Label Generator to analyze each diagnostic and will generate corresponding labels. The labels and other disruption-related information will be stored in MongoDB for further usage. (Reprinted from [26], Copyright (2020), with permission from Elsevier).

fusion feature extractor was first designed by combining CNNs and LSTMs together, namely the Fusion Feature Extractor (FFE), which is shown in figure 6.

CNNs were designed to extract temporal features with higher characteristic frequencies as well as spatial features containing profile information. Considering that each kind of

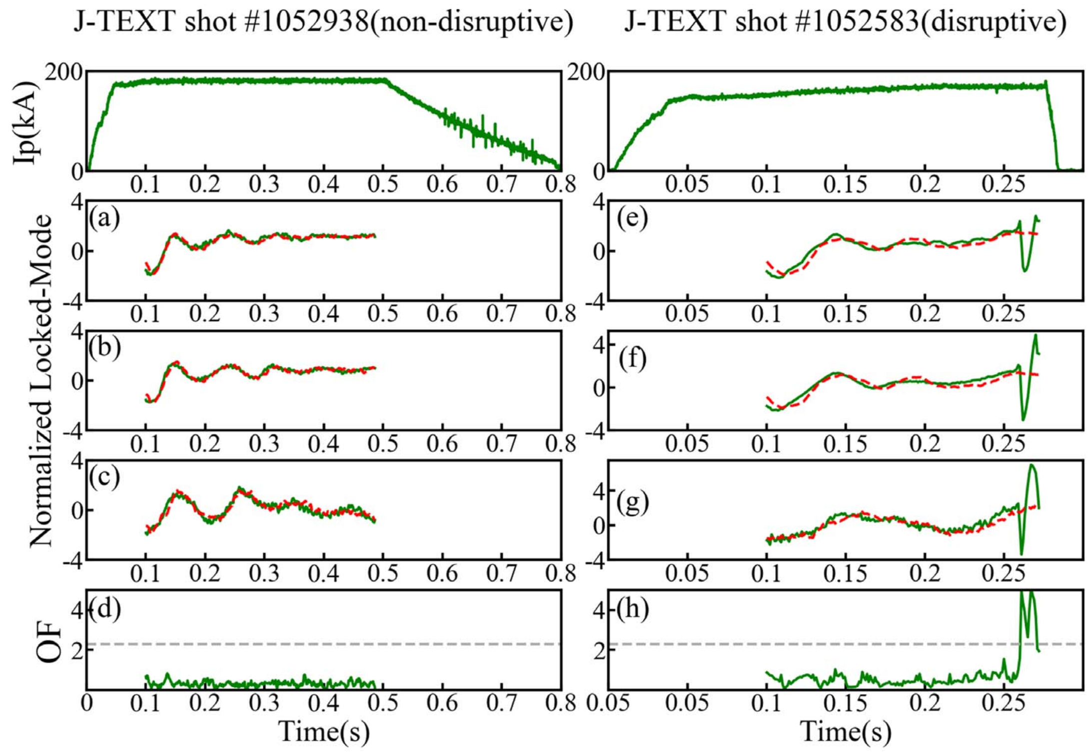

Figure 5. The predicted locked mode signal and the outlier factor of a typical non-disruptive and disruptive shot. The solid lines represent measured signals while the dashed lines represent predicted signals. Predictions of the non-disruptive shot #1052938 fit measured signals well and have a low outlier factor. Predictions of the disruptive shot #1052583 fit well at first, but worse when approaching the disruption. The outlier factor also improves simultaneously. (Reproduced from [27].

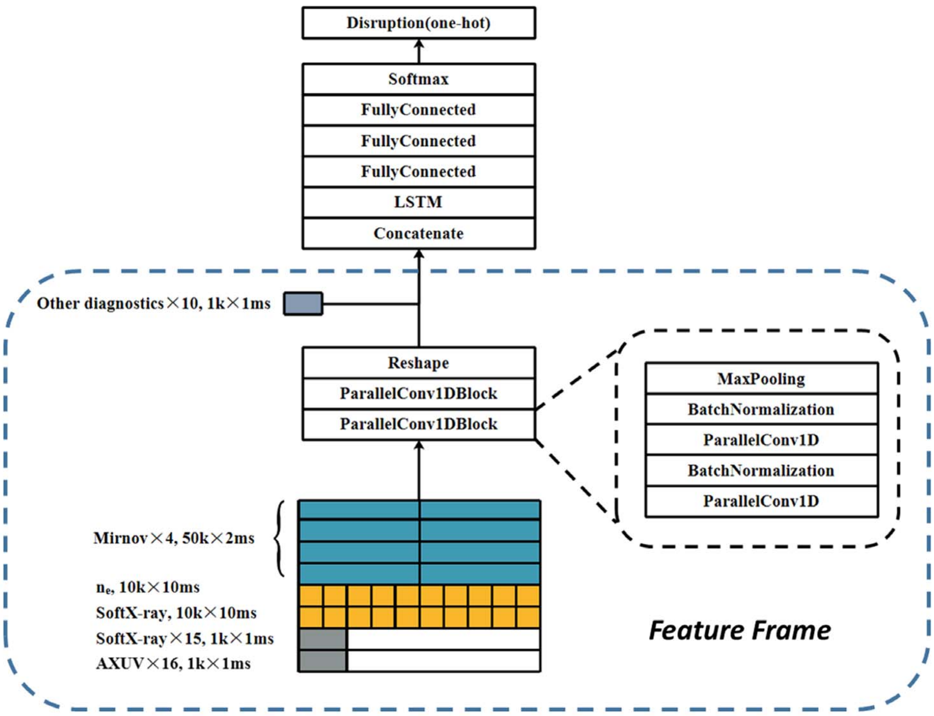

Figure 6. The structure of Fusion Feature Extractor-based disruption predictor. Inputs with high frequencies are processed to time windows with different sampling rates and time windows with different lengths and are fed into parallel 1D convolutional layers for feature extraction. The extracted features are concatenated with other inputs with lower frequencies and consist of a feature frame together. A sequence of the feature frame is then fed to the LSTMs and classifier to estimate whether the coming time step will be disruptive or not.

Table 1. Training data, strategy, and performance composition of all cross-tokamak experiments using pre-trained J-TEXT model and the EAST tokamak as the target tokamak to predict disruptions. The performance is measured by the AUC (Area Under the receiver operating characteristic Curve) value. Values in parentheses give the number of disruptive discharges.

Case No.	J-TEXT data	EAST data	Training strategy	Transfer strategy	AUC
1	None	1896 (355)	From scratch	/	0.821
2	None	20 (10)	From scratch	/	0.615
3	494 (189)	20 (10)	From scratch	/	0.661
4	None	None	Pre-trained	/	0.611
5	None	20 (10)	Pre-trained	Freeze & fine-tune	0.808
6	None	20 (10)	Pre-trained	Freeze & replace classifier	0.749

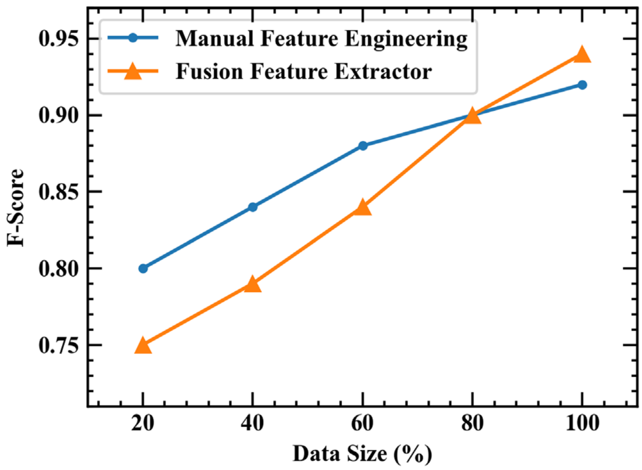

Figure 7. *F*-score from the test set of J-TEXT of disruption predictors based on SVM with feature engineering and FFE with different sizes of the training set. The performance of the FFE gradually improves and finally catches up with and outperforms that of the SVM with manual feature engineering as the training data accumulates.

diagnostic represents a different physical meaning, parallel 1D convolutional layers were applied so that the diagnostics with different meanings would not be mixed together. Furthermore, since different features bear different time scales and typical frequencies, different sampling rates and time-window lengths were applied for feature extraction to fit their own time scales. The features extracted by CNNs and other signals with larger time scales together consist of a feature frame, and were then fed into LSTMs for further feature extraction. Finally, the extracted features were sent to a fully-connected based disruption classifier to tell if the sample was disruptive or not.

The feature extractor was first tested on the J-TEXT tokamak, reaching a true positive rate of 96.36% and a false positive rate of 9.01%. Furthermore, a comparison between deep learning-based feature extraction and physics-guided manual feature extraction was made. Comparisons between different sizes of training sets applied to both methods were also compared, as shown in figure 7. With a small amount of training set fed to both models, manual feature engineering performed better than the deep learning method but was still not able to offer a satisfying result. As the data for training accumulated, the performance of both models gradually evolved, and the performance of the deep learning model finally overtook manual feature extraction.

The pre-trained model was then transferred to the EAST tokamak using a freeze fine-tune technique. The bottom layers that were designed to extract general features were frozen, while the top layers were device-specific and are supposed to be fine-tuned with data from the EAST tokamak. As a result, the transferred model trained with 10 disruptive and 10 non-disruptive discharges reached a similar performance to the model trained with 355 disruptive and 1541 non-disruptive discharges. Training data, strategy, and results are shown in table 1.

2.5. Interpretability research for disruption prediction models

Machine learning methods are usually considered as black-box methods, which means that they are so complex that they are not straightforwardly interpretable to humans in most cases. It is often the case that, machine learning methods are able to achieve promising results. However, the physical mechanisms are hidden between the parameters and can hardly be explained. Interpretability research for disruption prediction models may help to better understand how the models make decisions, and may also reveal new phenomena undiscovered before.

A new methodology for cross-machine study to find out what impacts the cross-machine predictor model building and to have a better understanding of the predictor was explored [29]. A new machine learning algorithm called LightGBM and a model interpretability tool SHAP (SHapley Additive exPlanations) were utilized to provide good performance prediction and enough analysis by applying to two different devices J-TEXT and HL-2A. The superiority of LightGBM is that it provides relative feature importance to analyze the differences between the two devices. To guarantee more detailed information on how the model made the prediction, the SHAP module was used to decompose each prediction into the contributions of each feature. A proposed general approach was to estimate the feature importance of two models which were trained from two different devices. For J-TEXT, the most important feature is the electron density, followed by soft x-ray (SXR) and radiated power, as shown in figure 8. On HL-2A, the most important feature is soft x-ray, followed by electron density and radiated power, as shown in figure 9. The models for the two devices were trained with

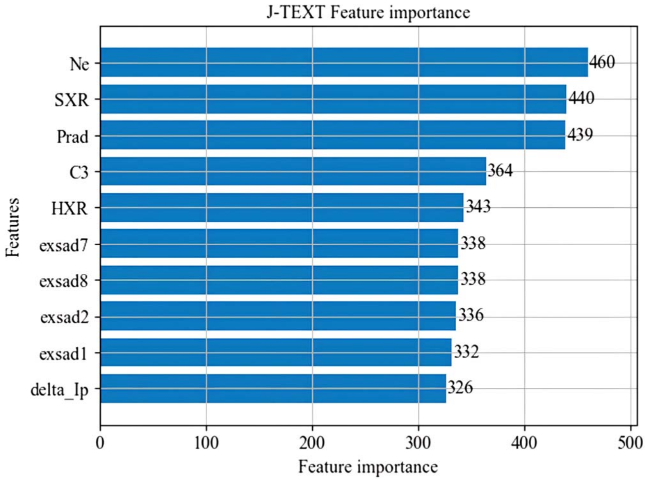

Figure 8. The feature importance given by SHAP on J-TEXT. The order of the signals represents their corresponding contributions, respectively. (Reproduced from [27].

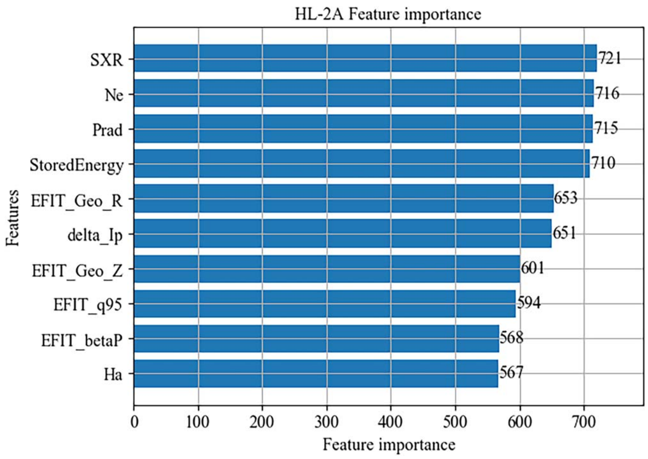

Figure 9. The feature importance given by SHAP on HL-2A. The order of the signals represents their corresponding contributions, respectively. (Reproduced from [27].

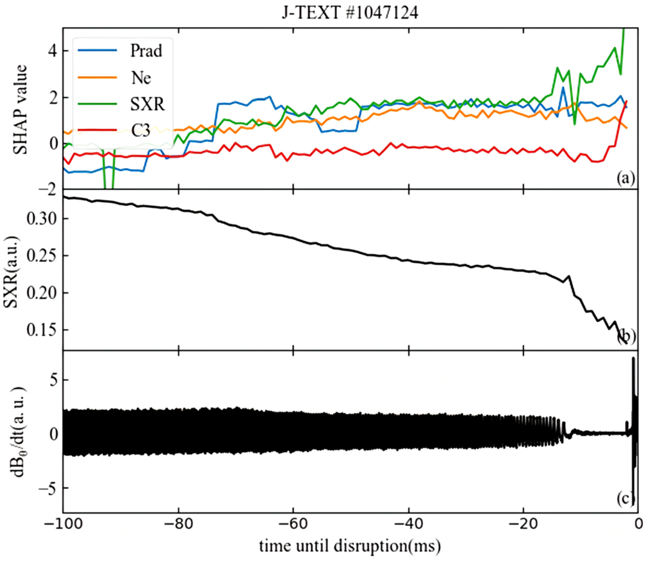

Figure 10. The SHAP value of 4 signals on J-TEXT varies over time and the precursor before disruption in 2 obvious signals: (b) soft x-ray radiation, (c) Mirnov signal. (Reproduced from [27].

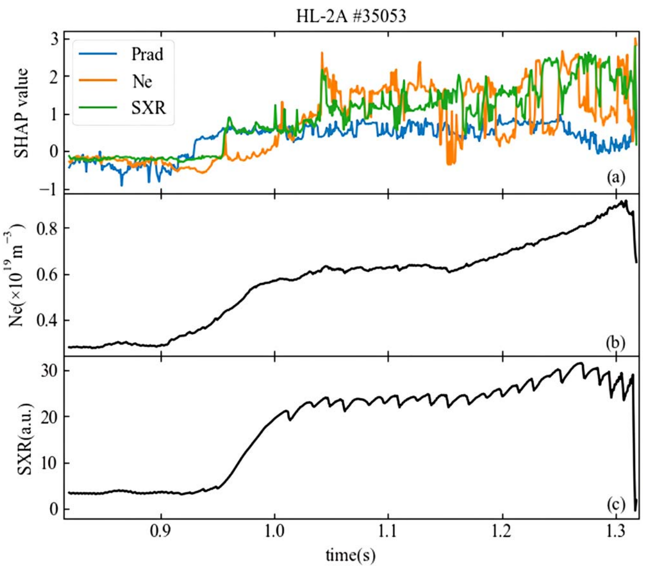

Figure 11. The SHAP value of 3 signals on HL-2A varies over time and the precursor before disruption in 2 obvious signals: (b) electron density, (c) soft x-ray radiation. (Reproduced from [27].

different sets of diagnostics because they did not have the exact same set of diagnostics. The models predicted disruptions well on both machines, respectively. However, the model had different judgments on the importance of features. The deep analysis of feature contribution for disruption was to find out the precursor of disruption, as shown in figures 10 and 11. When the disruption precursor approached, the models correctly extracted it and the feature contribution reflected this. The SHAP graph focused on the disruption discharges and the result may have some difference when compared with figures 8 and 9. Combined with these results, it can be confirmed that the disruption types in the two databases are different.

In addition to the aforementioned work, an interpretable disruption predictor based on physics-guided feature extraction called IDP-PGFE (Interpretable Disruption Predictor based on Physics-Guided Feature Extraction) was developed and applied on J-TEXT [30]. IDP-PGFE consists of a feature extractor, a disruption classifier, and an explainer, which is shown in figure 12. PGFE as a feature extractor extracts features with the inductive bias from the raw diagnostic signals based on the comprehension of phenomena. DART (Dropouts meet Multiple Additive Regression Trees) as a disruption classifier is a tree-based model using the trick of the dropout, which can reduce overfitting. SHAP as an explainer is an attribution approach, which uses a simpler explanation model as any interpretable approximation of the black-boxed model.

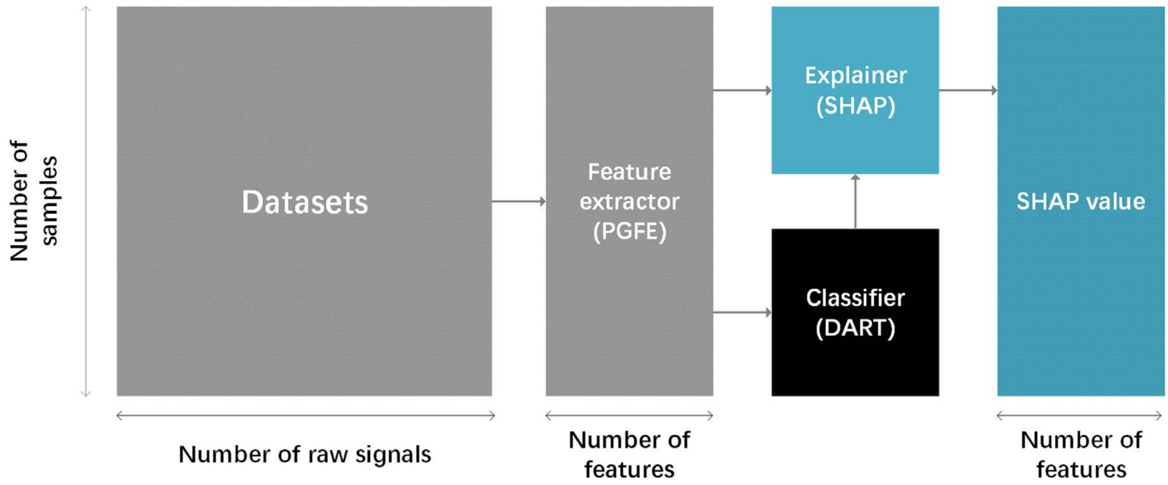

Figure 12. The overview of IDP-PGFE. IDP-PGFE consists of PGFE as a feature extractor, DART as a disruption classifier, and SHAP as an explainer. The explainer SHAP could interpret features extracted by PGFE. SHAP value represents the contribution of features.

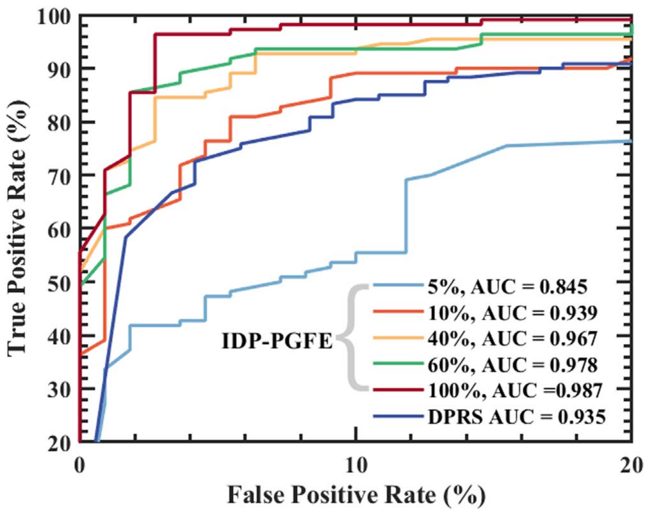

Figure 13. The ROC curves of IDP-PGFE trained by raw signals (DPRS - navy-blue) and different training sizes (100%-deep-red, 60%-green, 40%-yellow, 10%-orange, 5%-light-blue).

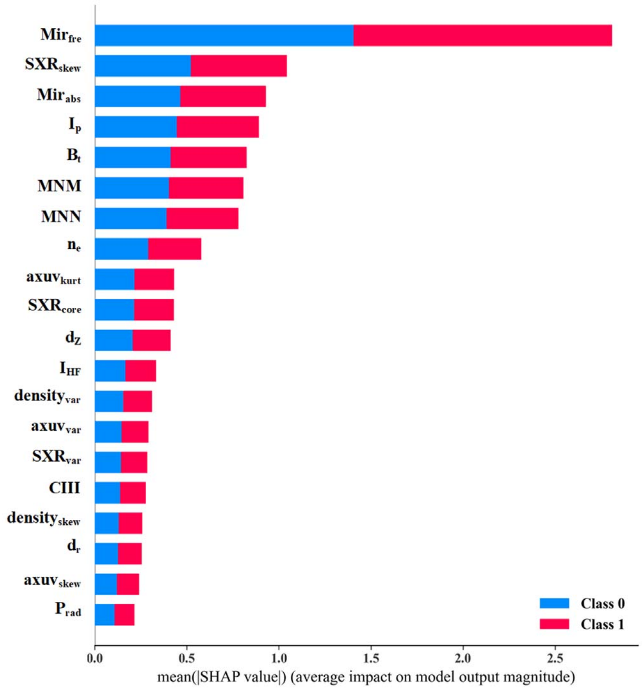

Figure 14. SHAP values of IDP-PGFE with extracted features. The order of the features represents their corresponding contributions, respectively. The mean of the SHAP value represents the average impact on model output magnitude. Class 1 represents the 'disruptive' impact on the model, while Class 0 represents the 'non-disruptive' impact on the model.

PGFE is the core component of IDP-PGFE, which makes the SHAP value more interpretable for tokamak disruption, improves the accuracy of the predictor, and lowers the data requirement of the predictor. IDP-PGFE reaches the best performance in J-TEXT disruption prediction tasks that are TPR (True Positive Rate) = 97.27/94.55/90%, FPR (False Positive Rate) = 5.45% with the tolerance of 10/20/30 ms, AUC = 0.98. PGFE could also reduce the data requirement of IDPOP. The ROC curves (Receiver Operating Characteristic curves) of IDP-PGFE trained by 100%, 40%, and 10% data size of the training set are shown in figure 13. The performance of IDP-PGFE using PGFE with only 10% data size of the training dataset is similar to the performance of disruption predictor using raw signals (DPRS) with a full training dataset.

The SHAP values of IDP-PGFE with extracted features are shown in figure 14. The most contributed feature in J-TEXT is the frequency of Mirnov probes, which reflects if there will appear locked mode. Because J-TEXT is a small-sized tokamak with a relatively smaller time scale, the locked mode amplitude is less important than the frequency of Mirnov probes. The interpretability was studied and also applied on J-TEXT density limit disruption (DLD) experiments. IDP-PGFE helps physicists to confirm that the application of ECRH did not lower the density limit, while the application of RMPs did raise the density limit. IDP-PGFE also has the potential for cross-machine disruption prediction and real-time disruption prediction.

3. Machine learning in the diagnostics data processing

Tokamak produces a large amount of diagnostic data each shot. A large portion of those data needs to be processed before revealing any physical meanings. Diagnostic data processing sometimes takes quite a lot of manual work. However, with the help of the power of machine learning, not only is the required manual work reduced but also extra results will be obtained compared with traditional methods.

The reliability of diagnostic systems in tokamak plasma is of great significance for physics research or even plasma control systems (PCS). Diagnostics signals are coupled with each other in tokamaks because the diagnostic systems are based on various principles and the sights of diagnostics are from various directions. There are times when the diagnostic signals are inaccurate because of the limitations imposed by the principles of the measurements. Given the situation, the measurements could be estimated or surrogated by other diagnostics with the help of machine learning methods. On J-TEXT, the electron temperature signals have been surrogated by SXR signals and other basic plasma signals, in case the diagnostic systems fail to detect plasma temperature. A fully-connected neural network was developed, utilizing back propagation with two hidden layers and generalized regression neural network [31]. The model was utilized to estimate plasma electron temperature approximately on the J-TEXT tokamak. The temperature profile was reconstructed by other diagnostic signals, and the average errors were within 5%.

Another application is to estimate horizontal displacement with an SXR array instead of using magnetic coils when RMPs are applied [32]. On J-TEXT, the plasma horizontal displacement is measured by magnetic coils. The coils will receive a large amount of interference if RMPs are applied, resulting in an inaccurate measurement. Although the interference by RMPs can be compensated with the RMPs coil current, the response of the plasma may introduce uncertainty to the compensated measurement. The SXR array can be used to estimate the plasma displacement as well. But instead of measuring the center of plasma current as the displacement coils do, the SXR array gives the center of radiation. There is a large gap between the measured radiation center and the plasma current center, and the gap between them is not constant and thus cannot be easily compensated. By training a neural network regression model that used the SXR array as input and magnetic horizontal displacement as the target, the trained model was able to accurately output the horizontal displacement by central current only using SXR diagnostics. In addition, it is worth noting that, in order to make the model more robust and be able to tolerate sensors' possible malfunction, all the samples had three randomly selected channels replaced by random noise during the training process. This made the trained model able to estimate the horizontal displacement with even up to 3 SXR channels malfunctioned with no visible performance degradation.

Electron Cyclotron Emission Imaging (ECEI) is a very powerful diagnostic. But the attenuation of the pixel channels must be tuned individually to the setup of the experiment. If the attenuation is not properly set up, the pixel channel will suffer from saturated or weak signals and will result in bad images. A module aimed at automatic data cleaning for J-TEXT ECEI signals based on machine learning was developed [33]. A 2-stage classifier model was designed and built, which was able to recognize six types of signal states: low-attenuation, saturated-background, high-background, weak signal, zero signal, and normal signal. By observing and analyzing the signals, three features were extracted manually. After the feature engineering, traditional machine learning algorithms like support vector machine (SVM) and random forest were used to train the classifiers. An accuracy of over 93% was achieved for classifying the aforementioned six signal states. The module could well indicate the problematic ECEI signals and could correct the channels with improper attenuation automatically, which significantly improved the quality of the image. The module could also provide the basis for further feedback regulation.

4. Real-time machine learning applications on J-TEXT

For the machine learning applications in the fusion experiment mentioned above, most of them need to be implemented in real-time. For example, the disruption prediction model needs to be run in real-time to trigger the disruption mitigation system, and the displacement estimation model needs to provide a real-time displacement value to PCS. The required control period for the applications above is around 1 ms, normally the same as the PCS control period. Different applications use different types of machine learning models. On J-TEXT, there are both traditional machine learning models and deep learning models just for disruption prediction. It is necessary to test and validate the feasibility of the real-time implementations of those models.

The first attempt was done during the density limit prediction and avoidance experiment. In the experiment, the disruption predictor needed to send a signal to the density feedback control system to shut the gas puffing valve as soon as it predicted a density limit disruption was coming. The inference interval of this model is set at 1 ms. As the model used here was pretty basic, although it was not a fully connected feedforward neural network, it could be coded as a series of matrices multiplication. This was implemented in LabVIEW-RT. The experiment was successfully conducted, and the density limit disruption could be avoided [25].

But for more sophisticated models, like neural networks constructed using the popular deep learning framework TensorFlow, a real-time application cannot be realized in the same way. On J-TEXT, we have developed a real-time control framework called J-TEXT Real-Time Framework (JRTF). It is based on real-time pre-empt patched Linux kernel. Using

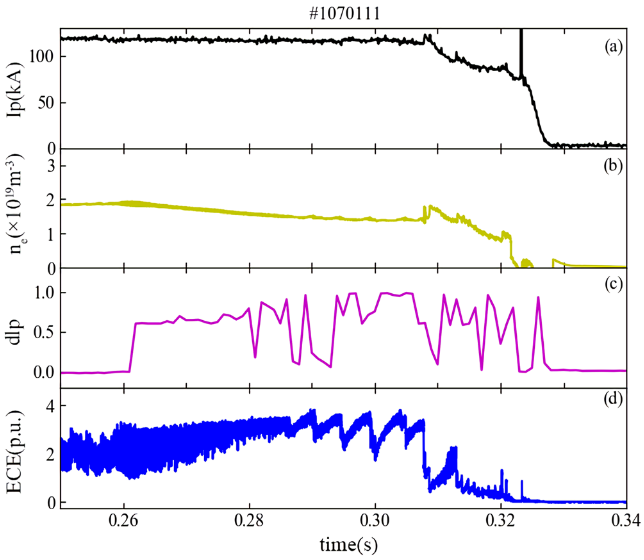

Figure 15. Density limit disruption without MGI. (a) I_p, (b) n_e, (c) disruption prediction and (d) ECE.

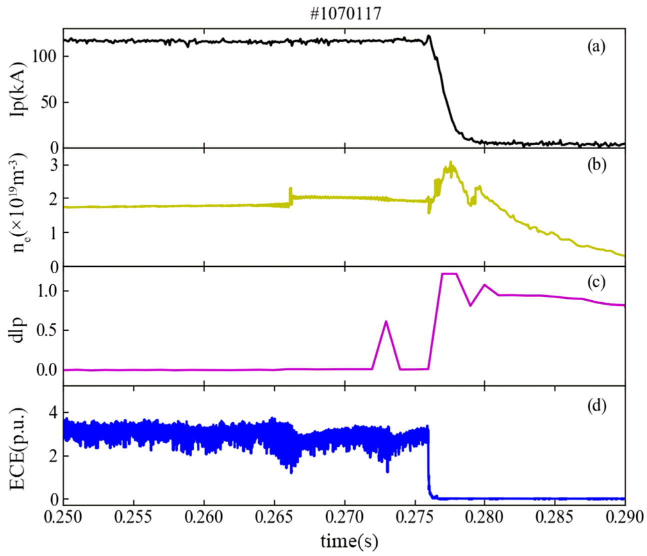

Figure 16. Density limit disruption with MGI before the disruption. (a) I_p, (b) n_e, (c) disruption prediction and (d) ECE.

JRTF, we have ported a TensorFlow constructed model, to be specific to the SXR horizontal displacement estimation model to a real-time environment. The model depends on the TensorFlow C++ library to run in real-time. The offline model is saved as a serialization format and is loaded by the C++ TensorFlow library. The model is able to finish one inference cycle in less than 200 μs. The system was tested in a piggyback experiment on J-TEXT [32]. The reason why we chose the horizontal estimation model to be implemented in real-time is that it was a deep neural network implemented by TensorFlow. Most of our deep learning models, including the disruption prediction models, are based on TensorFlow. The successful application of the model proves the feasibility of applying the TensorFlow-based model in real-time.

We also tried to implement a LightGBM-based traditional machine learning model using JRTF. It was connected to the trigger of the MGI value. The machine learning-based disruption predictor was applied to J-TEXT to confirm the reliability of integrating disruption prediction and the mitigation system. An exploratory experiment was proposed to estimate the performance of real-time disruption prediction. It was found that the disruption predictor was able to capture the precursors of density limit disruption and provide a warning time of 46 ms to prepare a mitigation system even if the MGI is offline, as shown in figure 15. When the MGI was prepared for the discharge, as shown in figure 16, the predictor successfully sent the predicted warning to it and only 4 ms was taken to shut down this discharge. The integrated disruption prediction and mitigation system proved to be effective when dealing with density limit disruption.

5. Summary

With the development of machine learning, its applications in physical experiments have taken off in recent years. On J-TEXT, machine learning-based research and applications in fusion plasma experiments have been one of the main subjects since 2013. With years of development, great results have been obtained in disruption prediction and diagnostic processing, as well as their applications in real-time experiments. Some of the models have been used on J-TEXT for rounds of campaigns to make joint decisions with the PCS and other systems to further support the dedicated experiments.

On J-TEXT, we first realized avoiding and mitigating density limit disruptions in real-time experiments based on a deep learning predictor. In addition, considering that future fusion reactors, such as ITER, can hardly bear disruption and thus cannot obtain enough disruptive discharges to train, we first proposed a disruption prediction model using only normal discharges. Furthermore, another attempt was carried out to transfer the pre-trained disruption prediction model to a larger target tokamak with limited disruption discharges. With respect to the interpretability of the machine learning-based models, an exploratory analysis was conducted to better reveal the decisions made by the models and the physical mechanisms of the disruption, as well as to discover new disruption-related phenomena. In addition, with years of development, the results achieved on J-TEXT have developed from initially predicting only one kind of disruption, such as locked mode or density limit-induced disruptions, to nowadays predicting all kinds of disruptions. The performance of the model has been developed from a true positive rate of ~80% and a false alarm rate of ~20% to a true positive rate of >90% and a false alarm rate of ~10%, also with a warning time of around 1 ms to over 30 ms.

Machine learning applications have been widely applied in fusion plasma experiments on J-TEXT tokamak. Promising results have been blooming in recent years. In the future, we will proceed to investigate machine learning-based techniques, to further address key problems in fusion plasma experiments, especially in disruption prediction, to better support the coming demands of ITER.

Acknowledgments

The authors are very grateful for the help of J-TEXT team. This work is supported by the National Key R&D Program of China (No. 2022YFE03040004) and National Natural Science Foundation of China (No. 51821005).

References

[1] Jumper J *et al* 2021 *Nature* **596** 583
[2] Bhuvaneswari V *et al* 2021 *Mater. Today Proc.* **46** 3263
[3] Baldi P, Sadowski P and Whiteson D 2014 *Nat. Commun.* **5** 4308
[4] Hernandez J V *et al* 1996 *Nucl. Fusion* **36** 1009
[5] Kates-Harbeck J, Svyatkovskiy A and Tang W 2019 *Nature* **568** 526
[6] Degrave J *et al* 2022 *Nature* **602** 414
[7] Mohapatra D, Subudhi B and Daniel R 2020 *Fusion Eng. Des.* **151** 111401
[8] Chang C S *et al* 2021 *Phys. Plasmas* **28** 22501
[9] Li H *et al* 2021 *Plasma Sci. Technol.* **23** 115102
[10] Li H *et al* 2022 *Nucl. Fusion* **62** 036014
[11] Li C *et al* 2018 *J. Instrum.* **13** P10029
[12] Hender T C *et al* 2007 *Nucl. Fusion* **47** S128
[13] ITER Physics Expert Group on Disruptions, Plasma Control and ITER Physics Basis Editors 1999 *Nucl. Fusion* **39** 2251
[14] de Vries P C *et al* 2011 *Nucl. Fusion* **51** 053018
[15] Boozer A H 2012 *Phys. Plasmas* **19** 058101
[16] Zakharov L E *et al* 2012 *Phys. Plasmas* **19** 055703
[17] Schuller F C 1995 *Plasma Phys. Control. Fusion* **37** A135
[18] Milani F 1998 Disruption prediction at JET *PhD Thesis* Aston University, Birmingham, UK
[19] Wroblewski D, Jahns G L and Leuer J A 1997 *Nucl. Fusion* **37** 725
[20] Yoshino R 2003 *Nucl. Fusion* **43** 1771
[21] Yang Q W *et al* 2006 *Chin. Phys. Lett.* **23** 891
[22] Ding Y H *et al* 2013 *Plasma Sci. Technol.* **15** 1154
[23] Wang Z X, Tang W K and Wei L 2022 *Plasma Sci. Technol.* **24** 033001
[24] Wang S Y *et al* 2016 *Plasma Phys. Control. Fusion* **58** 055014
[25] Zheng W *et al* 2018 *Nucl. Fusion* **58** 056016
[26] Zhang M *et al* 2020 *Fusion Eng. Des.* **160** 111981
[27] Zheng W *et al* 2020 *Plasma Phys. Control. Fusion* **62** 045012
[28] Zheng W *et al* 2022 arXiv:2208.09594
[29] Zhong Y *et al* 2021 *Plasma Phys. Control. Fusion* **63** 075008
[30] Shen C *et al* 2022 arXiv:2208.13197
[31] Dong J L *et al* 2021 *Plasma Sci. Technol.* **23** 085101
[32] Tang C Y *et al* 2022 *Fusion Eng. Des.* **182** 113248
[33] Zhang Z C *et al* 2022 *Fusion Eng. Des.* **177** 113065