Research on Recognition of Quench and Flux Jump Based on Machine Learning

Baobao Niu[1], Liangting Sun[2], Xianjin Ou[3]
[1]institute of modern physics, ChineseAcademy of Science.

Background

- The Fourth generation of Electron Cyclotron Resonance (**FECR**), which is currently being developed by the Institute of Modern Physics, uses **Nb3Sn** superconducting hexapole magnets with higher magnetic fields and composite structures.
- However, For Nb3Sn superconducting magnets, they exhibit significant thermal magnetic instability, known as flux jump. This characteristic will generate random voltage spikes during the excitation process of the magnet, leading to misjudgment by the Quench Detection System (QDS)
- To solve this problem, this study uses **machine learning algorithms** and aims to build a simplified and efficient recognition model to effectively distinguish the phenomenon of overshoot and flux jump during the excitation process of Nb3Sn magnets.

Methods and Results

Step 1: Construct the original data matrix

25 Flux jump

27 Quench

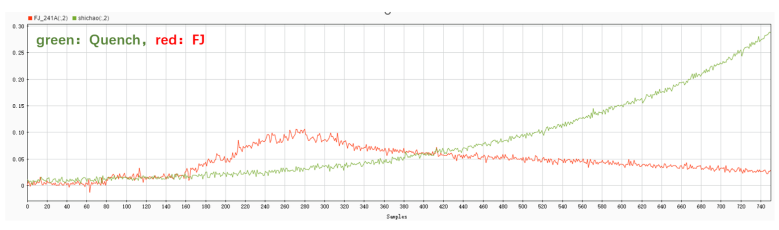


Fig1 : Comparison between Flux Jump and Quench

- 25 signals, 27 Quench signals
- Use the six-pole bridge voltage data
- Each data segment lasts for **150ms**, with a sampling interval of 0.2ms and a sampling frequency of **5000Hz**

Step2: Feature extraction

Frequency-domain feature	Definition
Absolute power	$P_A = log_{10}\left(\sum_{f=f1}^{f2} p_{xx}(f)\right)$
Relative power	$P_R = log_{10}\left(\sum_{f=f1}^{f2} p_{xx}(f) \Big/ \sum_{f=1}^{2500} p_{xx}(f)\right)$
Maximum power	$P_M = \log_{10} \max(P_{xx}(f)\|_{f1}^{f2})$
Center frequency	$c = f\|_{\log_{10}(\sum_{f=f1}^{f} p_{xx}(f))=\frac{1}{2}\log_{10}(\sum_{f=f1}^{f2} p_{xx}(f))}$

Time-domain feature	Definition
Standard deviation	$\sigma = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(x(t)-\mu)^2}$
The mean value of the first-order difference (MFD)	$\delta = \frac{1}{T-1}\sum_{t=1}^{T-1}\|x(t+1)-x(t)\|$
The mean value of second-order difference (MSD)	$\gamma = \frac{1}{T-2}\sum_{t=1}^{T-2}\|x(t+2)-x(t)\|$
The mean value of the standardized first-order difference (MSFD)	$\bar{\delta} = \frac{\delta}{\sigma}$
The mean value of the standardized second-order difference (MSSD)	$\bar{\gamma} = \frac{\gamma}{\sigma}$

Waveform feature	Definition
Skewness	$S = \frac{1}{T}\sum_{n=1}^{T}(\frac{x(n)-\mu}{\sigma})^3$
Kurtosis	$K = \frac{1}{T}\sum_{n=1}^{T}(\frac{x(n)-\mu}{\sigma})^4$
Activity	$A = var(x(n))$
Mobility	$M = \sqrt{\frac{var(diff(x(n)))}{var(x(n))}}$
Complexity	$C = \sqrt{\frac{Mob(diff(x(n)))}{Mob(x(n))}}$

- In this study, the 0-2500hz (single-sided frequency band) is divided into six frequency bands **(1-100)Hz, (100-300)Hz, (300-600)Hz, (600-1000)Hz, (1000-1500)Hz, and (1500-2500)Hz**, and four frequency domain features are extracted in each frequency band.
- The study found that this voltage signal is a non-stationary random signal, so its Frequency-domain feature need to be estimated using power spectrum estimation methods. Here, the **Welch method** is used, with a **blackman window** function as the window function.

Step 3: Construct the feature matrix

Colu	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Feature name	1-100Hz				100-300Hz				300-600Hz				600-1000Hz			
	Absolute power	Relative power	Maximum power	Center frequency	Absolute power	Relative power	Maximum power	Center frequency	Absolute power	Relative power	Maximum power	Center frequency	Absolute power	Relative power	Maximum power	Center frequency

Colu	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
Feature name	1000-1500Hz				1500-2500Hz				MFD	MSD	MSFD	MSSD	Activity	Mobility	Complexity	Kurtosis	Skewness
	Absolute power	Relative power	Maximum power	Center frequency	Absolute power	Relative power	Maximum power	Center frequency									

Tab1 : The selected feature and its number, name

- Given the strong correlation between standard deviation and Activity, only the Activity is retained.
- Set the FJ signal label to 1; set the quench signal label to 0;
- The resulting feature matrix is 52*34, where rows are samples (rows 1-25 are FJ, rows 26-52 are quench), and columns are features (columns 1-24 are frequency domain features, columns 25-28 are time domain features, columns 29-33 are waveform features, and column 34 is a label).

Step 4: Use classifier to classify

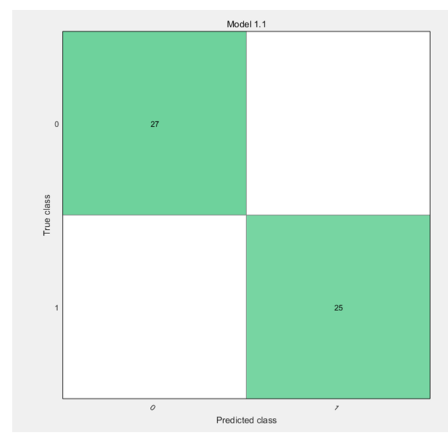
Model	kernel function	Accuracy
SVM	Linear	100%
	Quadratic	100%
	Cubic	100%
	Medium Gaussian	98.1%
KNN	Fine	96.2%
	Medium	88.5%
	Cosine	86.5%
	Weighted	92.3%
Complex Tree		73.1%
Logistic Regression		80.8%

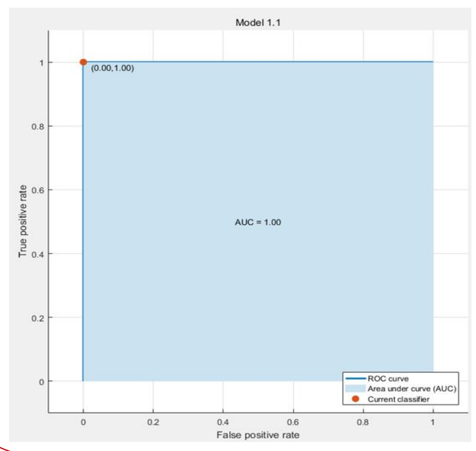
Tab2 : Accuracy of the original feature matrix on various classifiers

Model number 1.1
Status: Trained
Accuracy: 100.0%
Prediction speed: ~810 obs/sec
Training Time: 10.231秒

Classifier
Preset: Linear SVM
Kernel function: Linear
Kernel scale: Automatic
Box constraint level: 1
Multiclass method: One-vs-One
Standardize data: true

Feature Selection
All features used in the model, before PCA





- In this study, we used 10-folds cross-validation.
- Using SVM and the simplest linear kernel function, we can achieve 100% classification accuracy. The basic information of its model, confusion matrix, and ROC curve are shown in the figure above

Step 5: Construct the most concise and effective feature combination

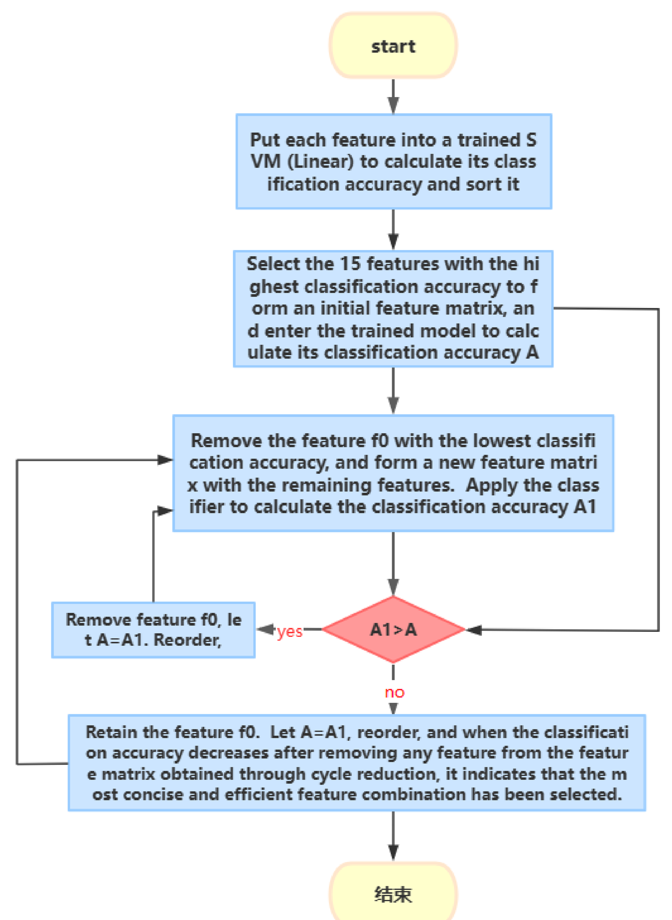


Fig2 : Flowchart of feature selection algorithm

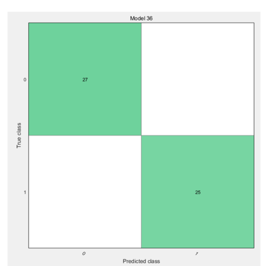
Number of feature	Name
21	1500-2500Hz Absolute power
31	Complexity
20	1000-1500Hz Center frequency
30	Mobility
5	100-300Hz Absolute power

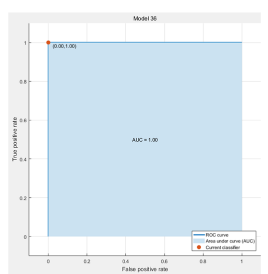
Tab3 : The optimal feature combination selected by the feature selection algorithm

Model number 36
Status: Trained
Accuracy: 100.0%
Prediction speed: ~3100 obs/sec
Training Time: 1.2731秒

Classifier
Preset: Linear SVM
Kernel function: Linear
Kernel scale: Automatic
Box constraint level: 1
Multiclass method: One-vs-One
Standardize data: true

Feature Selection
Used features, before PCA: column_5, column_20, column_21, column_30, column_31





Model	kernel function	Accuracy
SVM	Linear	100%
	Quadratic	98.1%
	Cubic	98.1%
	Medium Gaussian	100%
KNN	Fine	98.1%
	Medium	98.1%
	Cosine	98.1%
	Weighted	100%
Simple，medium，complex Tree		92.3%
Logistic Regression		100%

Tab4 : Accuracy of optimal feature combination on various classifiers

Number	feature selection algorithm	Feature combination	Classifier	Accuracy
1		1-33	linearSVM 10-folds cross validation	100%
2	my	5,20,21,30,31		100%
3	mRMR	20,21,23,31,33		96%
4	MulInf	8,21,23,31,32		96%
5	relief	5,8,21,23,31		94.2%
6	PCA			82.7%

Tab5 : Classification results of feature combinations obtained using different feature selection algorithms on the classifier

- Using the author's original feature selection algorithm (shown in Fig. 2), five optimal feature combinations were selected from 33 features (shown in Tab. 3), achieving a 100% classification accuracy on SVM (Linear).
- Then, the accuracy of the combination on different classifiers was compared (shown in Tab.4), verifying the excellent universality of these features.
- Furthermore, a comparison was made between common feature selection algorithms and the feature combinations selected by the author's feature selection algorithm, verifying the superior performance of the author's feature selection algorithm.

Conclusions

In this study, 27 quench samples and 25 flux jump samples were used, and 33 features were extracted from each sample. Multiple machine learning algorithms were used to train and build models on these data, and the accuracy of different algorithms was compared to explore the best recognition model. The experimental results showed that the model achieved a 100% classification accuracy on the linear kernel SVM using only 5 features. Using this machine learning model, high accuracy and computational speed were achieved in the identification of quench and flux jump, which can provide a reference for the optimization of subsequent FECR quench detection algorithms.