ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB345

MACHINE LEARNING WITH A HYBRID MODEL FOR MONITORING OF THE PROTECTION SYSTEMS OF THE LHC

C. Obermair*[1], M. Maciejewski, Z. Charifoulline, A. Apollonio, A. Verweij
CERN, Geneva, Switzerland
F. Pernkopf, Graz University of Technology, Graz, Austria
[1]also at Graz University of Technology, Graz, Austria

Abstract

The Large Hadron Collider (LHC) is the world's largest particle accelerator and uses a complex set of sophisticated and highly reliable machine protection systems to ensure a safe operation with high availability for particle physics production. The data gathered during several years of successful operation allow the use of data-driven methods to assist experts in finding anomalies in the behavior of those protection systems. In this paper, we derive a model that can extend the existing signal monitoring applications for the LHC protection systems with machine learning. Our hybrid model combines an existing threshold-based system with a Support Vector Machine (SVM) by using signals, manually validated by experts. Even with a limited amount of data, the SVM learns to integrate the expert knowledge and contributes to a better classification of safety critical signals. Using this approach, we analyze historical signals of quench heaters, which are an important part of the quench protection system for superconducting magnets. Particularly, it is possible to incorporate expert decisions into the classification process and to improve the failure detection rate of the existing quench heater discharge analysis tool.

INTRODUCTION

The early detection of faulty components contributes significantly to increasing machine availability and, thus, the LHC's physics potential expressed in terms of integrated luminosity. To protect the highly critical systems the machine protection system ensures safe operation of accelerator equipment (e.g. the superconducting magnets) and protects it from damage [1]. If the system fails it can result in an LHC downtime in the order of three months. Therefore, it requires consistent supervision of the components through signal monitoring and regular hardware commissioning tests. The quench protection system is part of the machine protection system and the Quench Heaters (QHs) are an essential part of it. The purpose of the QHs is to expand the quenching region of a superconductor, in order to enlarge the area of energy dissipation and, thus, reduce the potentially dangerous hot spot temperatures in the superconducting material. All of the 1232 LHC main dipole magnets are equipped with eight QH circuits. During magnet operation four out of the eight QH circuits are ready to be operated in case of a magnet quench. The other four QH circuits provide redundancy, in case of a fault in one of the other four circuits.

* christoph.obermair@cern.ch

Existing signal monitoring applications are based on the calculation of characteristic features representing a signal, which are compared to fixed thresholds. These thresholds give a clear answer whether the signal is healthy or faulty. This approach is particularly effective because experts can incorporate their knowledge about the behavior of a component's degradation into the analysis process. The quench heater discharge analysis tool [2] is one application which makes use of this, but several other components of the LHC machine protection system use the same approach [3–5].

Since the first commissioning of the LHC in 2008, the amount of system supervision data is growing and alternative signal monitoring approaches such as machine learning are currently gaining attention [6]. Several efforts have been made in the past to show the potential of machine learning for LHC protection systems, e.g. to observe anomalous behaviors of LHC superconducting magnets [7]. Another approach [2] used a feed forward neural network to analyze quench heater discharges. However, the lack of "faulty" signals often prevents machine learning models to reach the necessary reliability to replace existing analysis tools [2].

In other fields of research [8–10] it is common to build ensembles of different classifiers in order to make use of the so called *wisdom of the crowd* [11], which allows classifiers to contribute to a better overall classification result (e.g. XGBoost [12]). However, a hybrid approach which combines the advantages of existing LHC signal monitoring applications with the advantages of machine learning models has not been considered yet. Thus, the objective of this work was to develop an approach that allows LHC signal monitoring applications to benefit from the growing amount of historical data.

The paper is structured as follows: First the concept of threshold-based signal monitoring applications is explained and the workflow of the hybrid classification approach is derived. Subsequently, the existing analysis of past quench heater discharges is presented and the results of the new hybrid classification approach are discussed. Finally, the strengths and limitations of the approach are discussed and the conclusion is presented.

DEVELOPMENT OF THE MODEL

Typically, an LHC signal monitoring application processes windows of C multivariate discrete signals. Those data batches are represented by $Z_d \in \mathbb{R}^{N \times C}$, where N is the amount of samples and $d \in [1, ..., D]$ is the window index. Depending on the use case, such a batch can either have a

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB345

fixed or variable length and the beginning is either defined by a specific event, like a quench, or a manually chosen event, like a particular state of operation represented by the so-called beam mode.

Threshold-Based Classification

The workflow of a threshold-based signal monitoring approach is summarized in Fig. 1. Formally this means, that out of each signal batch matrix Z_d a function ϕ_f calculates F features $x_d \in \mathbb{R}^F$.With those features x_d, the application then assigns a label "healthy" ($y_d = 1$) or "faulty" ($y_d = -1$) to each batch with a threshold function:

$$g(x_d) = \begin{cases} 1 & \text{if } \check{k} < x_d < \hat{k} \\ -1 & \text{otherwise,} \end{cases} \tag{1}$$

in which $\check{k}$ is the minimum threshold vector and $\hat{k}$ is the maximum threshold vector, both determined by experts.

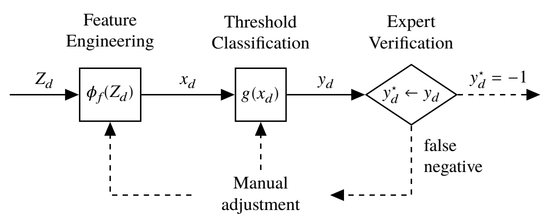

Figure 1: Workflow of a threshold-based signal monitoring approach. First the features are calculated, then a threshold is set to assign a label to the signal. This label is validated by experts in the last step.

In case the signal condition is identified as "faulty", experts have to verify this result. Consequently, if the experts decide that the prediction of the classification algorithm was true negative ($y_d^\star = -1$), they can initiate further actions, like a hardware inspection. However, if the experts decide that the automatic signal classification does not reflect the actual condition of a component (false negative), the machine operation continues as usual. Furthermore, the experts could then adjust the classification algorithm and/or the thresholds, such that this specific "faulty" classification does not occur in the future. However, due to the high amount of signals, experts often only get notified in case of a "faulty" classification. This means they can only intervene if the classification was false negative. A false positive label, which is a "faulty" signal labeled as "healthy", only emerges if other protection systems are triggered or if damage occurs.

Machine Learning Based Classification

During classification with machine learning models the parameters of a threshold function are optimized such that the best classification on a given input data set is reached. Common classification algorithms include logistic regression, random forest, neural networks, and SVMs [13]. We use the latter as it is especially suited for handling data sets with limited amount of samples and high dimensions.

The workflow of a machine learning based classification is similar to the threshold-based classification, but the threshold function is defined by the separation hyperplane:

$$h(x_d) = w^{\mathrm{T}}\phi(x_d) + b, \tag{2}$$

where w contains the weight parameters, b is the bias parameter and $\phi(x_d)$ is a fixed feature space transformation. Those parameters are determined by solving the following optimization problem, using training data, i.e.

$$\begin{aligned} &\arg\min_{w} \quad \left(\frac{1}{2}\|w\|^2 + C\sum_{d=1}^{D}\xi_d\right) \\ &\text{subject to: } y_d^\star(w^{\mathrm{T}}x_d + b) \geq 1 - \xi_d, \quad d = 1,...,D \\ &\qquad \xi_d \geq 0, \end{aligned} \tag{3}$$

where ξ_d is a slack variable for soft classification which handles misclassified data samples or anomalies in the data set, and C is a parameter that determines the importance of the outliers. Furthermore, the radial basis function is chosen as a kernel $\phi(x_d)^{\mathrm{T}}\phi(x_d')$ [13].

Hybrid Classification

The hybrid classification extends the threshold-based signal monitoring with machine learning, such that the amount of false negative classified labels is minimized and less manual adjustments by experts are necessary. The workflow of the hybrid classification approach is shown in Fig. 2. For systems with a high repetition rate it is important to make threshold adjustments automatically. In the hybrid classification approach this adjustment is handled by an SVM.

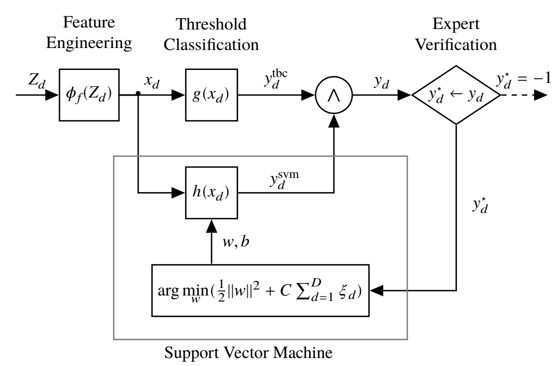

Figure 2: Hybrid classification approach. Similarly to the threshold based approach, first the features are calculated, then a label is assigned to the signal, which is consequently checked by experts. The SVM learns from the past decisions of the expert, by optimizing the parameter w in the separation hyperplane $h(x_d)$.

Specifically, an SVM performs the continuous threshold adjustment, while the threshold-based signal monitoring application operates with fixed thresholds from experts. In the initial phase, the available historical data is used to determine the parameters w and b of the SVM separation hyperplane $h(x_d)$. For each new batch d, the threshold-based classification provides y_d^{tbc} and the SVM classification determines y_d^{svm} from x_d. The output y_d is then determined by combining both outputs with a logical AND, i.e,

$$y_d = \begin{cases} 1 & \text{if } y_d^{\mathrm{tbc}} = 1 \wedge y_d^{\mathrm{svm}} = 1 \\ -1 & \text{otherwise,} \end{cases} \tag{4}$$

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB345

which is verified by experts. Once the experts have evaluated the condition of the component, the corresponding label is added to the training set and the SVM parameters w and b are recalculated. Accordingly, the new label is used as feedback for future decisions of the SVM.

APPLICATION OF THE MODEL

In this section the previously applied hybrid model is applied to the classification of QH breakdowns in the main dipoles of the LHC. The QH discharges are currently analyzed by the Quench Heater Discharge Analysis (QHDA) tool, which groups QH discharges into “healthy” and “faulty” with a threshold-based classification system. An extensive analysis takes place following each quench event before the main dipole can be powered again. In case of a quench in one of the main dipoles, the experts have several hours to check a “faulty” classification of a QH discharge before the magnets can be powered again, which is why a false negative classification (damage predicted while no damage) has a limited impact on the availability of the LHC. A false positive classification (no damage predicted while damage) has to be avoided by all means.

The QHDA tool validates the QH discharges using the following features [2]:

1. *Steady state voltage level:* The initial and final values of the voltage are used.
2. *Characteristic time of the pseudo-exponential decay:* The characteristic time of the pseudo-exponential decay is determined from the voltage and the current signals during the QH discharge.
3. *Steady state resistance level:* The initial resistance of the QH strip is determined from the voltage and current signals.
4. *Signal comparison:* The signals and the above features are compared sample-wise to the reference discharge of the corresponding QH circuit.

Some failures and precursors of failures in the QH circuits are difficult to detect with a threshold based method because they might correlate with other characteristics, which are not verified or which are sensitive to case by case variations. For example the initial resistance of the QH strip is calculated from the voltage and current signals at the start of the QH discharge. Due to the properties of the QH circuit there can be a switch-on delay, oscillations, and noise in both signals, which can cause variation in the calculated initial resistance from discharge to discharge.

Therefore, it was studied, whether a hybrid classification could identify such correlations and if it can consequently decrease the amount of false positive classifications.

RESULTS

The QHDA is implemented into an environment called *LABView*. As machine learning algorithms are commonly implemented in *python*, the features of the QHDA tool were reimplemented in the environment of the “LHC Signal Monitoring Project” [14] to recreate the discussed approaches. The values of the fixed thresholds have been set by experts and the hyperparameters of the SVM have been optimized using training data. The hybrid approach is implemented as stated before, i.e. combining the threshold-based (TB) classification with the SVM classification. The data-set contains stored discharges from 2014 to 2018. 3130 QH discharges have been labeled as “healthy” and come from 1230 main dipole magnets. 116 discharges were labeled as “faulty” and come from 68 different dipole magnets. This data-set was labeled by experts, who classified each discharge, marking even small deviations as “faulty”.

Table 1 compares the different methods by their performance. The true positive (TP) rate is the fraction of correctly identified “healthy” discharges relative to the total amount of “healthy” discharges. On the other hand the false positive (FP) rate defines the amount of falsely labeled “healthy” discharges relative to all “faulty” labeled discharges. The TB model indicates the rebuilt QHDA tool in python.

Table 1: Results of Different Performance Measures

Method	TP rate	FP rate
QHDA model	0.999	0.147
TB model (rebuilt QHDA)	0.993	0.078
SVM model	0.991	0.131
Hybrid model	0.989	0.024

From Table 1 it can be seen, that the TB model differs from the QHDA, due to slight variations in the calculation methods, but they both have a relatively good TP rate and a relatively bad FP rate. The hybrid model demonstrates a significantly improved FP rate, while it shows a small degradation in the TP rate. The FP rate of 14.7% for QHDA does not mean that 14.7% of the recorded discharges caused damage to the QHs or the magnets, but indicates the fraction of cases, which required further investigations by an expert. Consequently, experts were only missing 2.4% of the cases, which the hybrid model classified as “faulty”. This shows that machine learning improves the identification of cases, which experts need to investigate compared to simpler TB algorithms. Furthermore, ML with an SVM will remember previous expert decisions.

CONCLUSION

In this paper a promising concept for complementing traditional threshold-based LHC signal monitoring tools with machine learning is presented, illustrated by the example of QH signal analysis. This is achieved by building an ensemble of the existing signal monitoring application and an SVM, which is trained on historical data. The conducted analysis showed that this hybrid classification approach reduced the FP rate of the existing QHDA tool from 14.7% to 2.4%. Furthermore, the new approach allows the automatic incorporation of expert decisions into the classification process.

Overall, it has been demonstrated that even a limited amount of historical data can be beneficial for signal monitoring applications through the support of machine learning.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB345

REFERENCES

[1] R. Schmidt *et al.*, “LHC machine protection,” in *Proc. 22nd Particle Accelerator Conf. (PAC’07)*, Albuquerque, NM, USA, Jun. 2007, paper TUZAC03, pp. 878–882.

[2] Z. Charifoulline *et al.*, “Overview of the Performance of Quench Heaters for High-Current LHC Superconducting Magnets,” *IEEE Transactions on Applied Superconductivity*, vol. 27, no. 4, pp. 1–5, 2017, doi:10.1109/TASC.2016.2642991

[3] Z. Charifoulline *et al.*, “Resistance of Splices in the LHC Main Superconducting Magnet Circuits at 1.9 K,” *IEEE Transactions on Applied Superconductivity*, vol. 28, no. 3, pp. 1–5, 2018, doi:10.1109/TASC.2017.2784355

[4] E. Holzer *et al.*, “Beam Loss Monitoring System for the LHC,” in *Proc. IEEE Nuclear Science Symposium Conference Record*, Fajardo, PR, USA, 2005, pp. 1052–1056, doi:10.1109/NSSMIC.2005.1596433

[5] J. Wenninger, “Machine Protection and Operation for LHC,” *Proceedings of the Joint International Accelerator School: Beam Loss and Accelerator Protection*, vol. 2, no. 1, pp. 377–401, 2016, doi:10.5170/CERN-2016-002.377

[6] U. Gentile and L. Serio, “A Machine-Learning Based Methodology for Performance Analysis in Particles Accelerator Facilities,” in *2017 European Conference on Electrical Engineering and Computer Science (EECS)*, Bern, Switzerland, 2017, pp. 90–95, doi:10.1109/EECS.2017.26

[7] M. Wielgosz, A. Skoczeń, and M. Mertik, “Using LSTM recurrent neural networks for monitoring the LHC superconducting magnets,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 867, no. 1, pp. 40–50, 2017, doi:10.1016/j.nima.2017.06.020

[8] A. Kharrat, K. Gasmi, M. B. Messaoud, and M. Abid, “A Hybrid Approach for Automatic Classification of Brain MRI Using Genetic Algorithm and Support Vector Machine,” *Leonardo journal of sciences*, vol. 17, no. 1, pp. 71–82, 2010.

[9] M. Seera, C. P. Lim, S. Nahavandi, and C. K. Loo, “Condition monitoring of induction motors: A review and an application of an ensemble of hybrid intelligent models,” *Expert Systems with Applications*, vol. 41, no. 10, pp. 4891–4903, 2014, doi:10.1016/j.eswa.2014.02.028

[10] L. Rokach, “Ensemble-based classifiers,” *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010, doi:10.1007/s10462-009-9124-7

[11] J. Surowiecki, *The wisdom of crowds*, 1. ed. New York, NY: Anchor Books, 2005.

[12] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi:10.1145/2939672.2939785

[13] C. M. Bishop, *Pattern recognition and machine learning*. New York, NY, USA: Springer, 2006.

[14] M. Maciejewski, Z. Charifoulline, A. Verweij, and P. Hagen, Signal Monitoring, https://sigmon.web.cern.ch/about.