ISBN: 978-3-95450-184-7 doi:10.18429/JACoW-IPAC2018-WEPAF062

MACHINE LEARNING METHODS FOR OPTICS MEASUREMENTS AND CORRECTIONS AT LHC

E. Fol*, F. Carlier, J. Coello de Portugal, A. Garcia-Tabares Valdivieso, R. Tomás
CERN, Geneva, Switzerland

Abstract

The application of machine learning methods and concepts of artificial intelligence can be found in various industry and scientific branches. In Accelerator Physics the machine learning approach has not yet found a wide application. This paper is devoted to the evaluation of machine learning methods aiming to improve the optics measurements and corrections processes at LHC. The main subjects of the study are the detection of faulty beam position monitors and the prediction of quadrupole errors using decision trees and artificial neural networks. The results presented in this paper clearly show the suitability of machine learning methods for the optics control at LHC and the potential for further investigation on appropriate approaches.

INTRODUCTION

With the increased technological complexity of accelerators, meeting the demand of accelerator control and operation necessitates more powerful and faster methods. Machine learning methods and concepts of artificial intelligence are considered in various industry and scientific branches, and recently, these methods have been used in high energy physics mainly for experiments data analysis [1, 2].

The central task in the optics measurements and corrections processes is the identification of optics imperfections and computation of corrections. In this work, we present a machine learning approach for corrections computation and comparison of prototyped models trained on simulations data.

A problem appearing often in the measurement process is the identification of faulty data samples in measurements that requires application of automatic tools as well as human intervention. The application of cluster analysis [3] to detect faulty monitor data in early stages of optics measurements and corrections process is also a part of the presented work.

The objective of this preliminary study is to investigate the effectiveness of machine learning methods applied to accelerator optimization, accelerator control and in particular on optics measurements and corrections. Presented experiments on simulated and real data provide the basis for further studies such as application of more complex neural network structures, combination of different noise detection algorithms and integration of these methods into real-time machine operation.

* elena.fol@cern.ch

FAULTY BPMS DETECTION USING CLUSTER ANALYSIS

The phase of measured betatron oscillation is inferred from a harmonic analysis of the turn-by-turn transversal beam positions measured at BPMs around the ring. Reconstruction of the optics is mainly concerned by the propagation of phase advances between BPMs [4]. Therefore, the appearance of a faulty signal has significant impact on the obtained optics functions and sequentially computed corrections. In this work we evaluate faulty BPM signal detection applying cluster analysis. Cluster analysis includes methods of grouping or separating data objects into regions in a hyperparameter space, such that dissimilarity between the objects within each cluster is smaller than between the objects assigned to different clusters.

Motivation

The main issue regarding the problem of faulty BPMs is the appearance of unphysical data in the reconstructed optics functions. Currently used data cleaning techniques are based on identification of flat signal, threshold definition for the signal spikes and manual cleaning of the data [5]. Most of the noise can be removed using these methods, as well as through applying advanced signal improvements techniques based on SVD and FFT [6], however, faulty data samples can be observed in the optics functions.

The fact that the properties of bad signal producing unphysical data in the reconstructed optics functions are often unknown requires alternative solutions to detect faulty BPMs. The application of cluster analysis should replace the basic numerical tools currently used to identify the outliers in the measured data and thus reduce the human intervention in the data analysis process. The signal improvements techniques should remain a part of the signal processing.

Conceptual Solution

Giving the described constraints, clustering appears as an appropriate alternative method since the analysis can be performed on multi-dimensional space including all parameters of turn-by-turn data obtained through the harmonic analysis [7]. Another benefit of cluster analysis is the *unsupervised learning* approach, which means that no labeled data is needed to train the algorithm. This property of cluster analysis has significant importance since we do not aim to replicate the results of existing techniques and no training data set is available.

The analysis is performed on a three dimensional parameter space containing the betatron tune, the amplitude of the measured oscillations and the relative difference between the

ISBN: 978-3-95450-184-7 doi:10.18429/JACoW-IPAC2018-WEPAF062

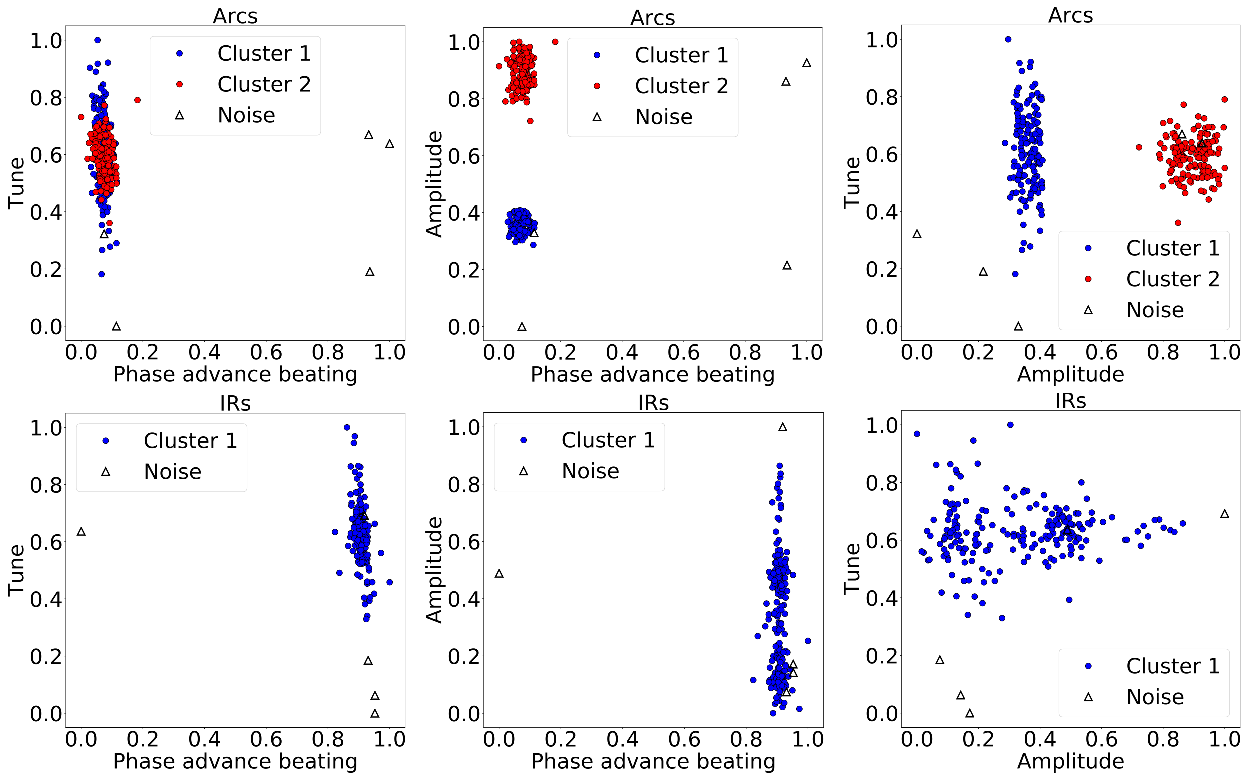

Figure 1: 2D-projection of the clusters produced by DBSCAN. The data is scaled to the range [0,1].

measured and design phase advances between the BPMs.The measurements from the interaction regions (IRs) and the arcs are treated separately since the data points distribution is different in these regions. The cluster analysis presented in this section was applied on a data measurement set taken during the machine development using half integer tunes at injection energy.

Defining the faulty signal recognition problem as a clustering task, the data has to be separated into minimum two clusters - good and faulty signal, which can be identified as outliers or considered as noise. Since the appearance of outliers in the data affects the computation of the mean of parameters, the algorithms based on centroids search such as K-means [8] are not appropriate for our problem. Instead of centroid search, the clusters can be built based on the density of the data samples.

Results Applying DBSCAN

The DBSCAN (Density-based spatial clustering of applications with noise) [9] is a data clustering algorithm for large spatial databases. The algorithm assumes clusters of arbitrary shape and views clusters as areas of high density separated by areas of low density, instead of looking for the centroids. The main idea of the algorithm is the concept of core samples, which are samples in areas of high density. A cluster is therefore a group of core samples which have to contain a minimum number of points in the neighborhood, which is computed by the chosen distance metric. The cluster also includes non-core samples that are in the neighborhood of a core sample, but are not core samples themselves (as there are not enough points in the neighborhood). Consequently, the input of the algorithm is the neighborhood distance *Eps* and the minimum number of data points *MinPts* in a neighborhood of a core point.

Depending on the tuning of input parameters of the DBSCAN algorithm, one or more clusters can be built. To indicate bad BPMs we can consider one cluster, the data points outside the cluster should be considered as signal produced by faulty BPMs. Alternatively, building several clusters of good BPMs is potentially interesting to discover the properties of good BPMs and the relations between the measured parameters. The study of common properties of the BPMs within each of the produced clusters can help to extract new features of the measured data.

The results of applying DBSCAN with *Eps*=0.3, *MinPts*=60 and *Eps*=0.5, *MinPts*=30 respectively for the arcs and IRs measurements are presented in Fig. 1. Different definitions of minimum distance is required since the variance of the data in the IRs is significantly higher compared to the arcs. This difference is due to the large β-functions in the IRs with small IP beam size. Moreover, within the arcs a data separation is also expected since the optics inside the arcs is periodic due to the presence of alternating focusing and defocussing quadrupoles. The obtained clustering constellation in the arcs is expected since it corresponds to two different lattice regions - focusing and defocussing areas. The IRs measurement build only one cluster, which allows to identify the outliers.

ISBN: 978-3-95450-184-7 doi:10.18429/JACoW-IPAC2018-WEPAF062

To evaluate the effectiveness of cluster analysis the reconstructed optics has to be studied. We use beta-beating as the most relevant optics control parameter since it describes the difference between the current optics and the designed machine. Figure 2 shows the beta-beating measurement obtained from optics analysis performed on turn-by-turn data excluding the BPMs identified by DBSCAN as noise. The achieved elimination of significant number of unphysical data samples demonstrates the potential of DBSCAN on the detection of faulty BPM signal.

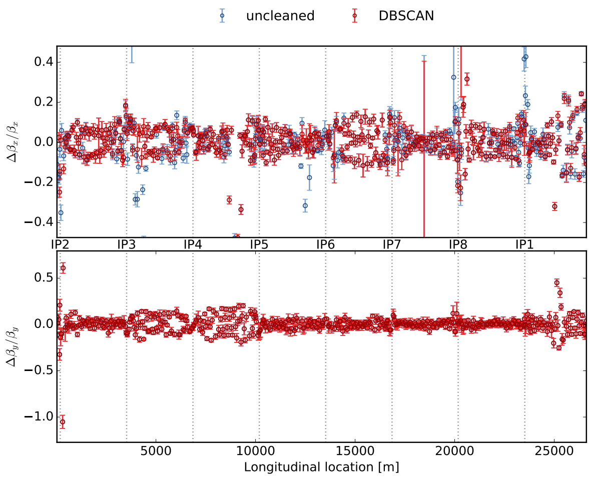

Figure 2: Comparison of beta-beating reconstructed from uncleaned and DBSCAN cleaned data.

Outlook

As the DBSCAN algorithm is robust against the outliers, it can be applied to eliminate them on different stages of measurements and correction process. Although different input parameters for the algorithm might be required for specific applications.

The clustering methods allow to analyze a space of arbitrary dimensionality and thus an arbitrary number of optics parameters can be selected to perform the analysis. This allows to observe the relation between particular observables and the patterns in the clustered data. The fact that particular good BPM data points can be separated in different clusters depending on the algorithm settings is of interest for further investigation. A deeper analysis of clustering patterns could cast light on unobserved properties of BPMs.

QUADRUPOLE ERRORS PREDICTION

To compute the corrections, the measured data has to be compared with the design machine. The deviations from design have to be identified by special analysis methods and compensated by improved machine settings according to computed corrections [4, 10–13]. In terms of machine learning, the corrections computation can be defined as a regression problem that can be solved by training a model using past measurements and corresponding corrections [14].

Considering the problem of predicting the quadrupole errors, *supervised learning* approach applying a set of regression models is used. Given a set of features $X = x_1, x_2, ..., x_m$ and target y, an estimator can learn a regression model or a non-linear approximation. As regression methods we use Random Forest decision trees [15], Orthogonal Matching Pursuit (OMP) [16] and Multi-layer Neural Network [17]. The regression models are trained and validated on two different datasets for nominal β^*=40 cm and injection optics used in 2016. As input parameter we use the differences between nominal model and simulated perturbed optics. The regression models are trained to predict the quadrupole errors which produce the perturbations in the input data. To validate the performance of different methods shown in Table 1 we use the Mean Absolute Error (MAE) and the explained variance as figures of merit.

Table 1: Performance of Applied Models

Injection optics

Model	**MAE** $[10^{-5}m^{-2}]$	**Explained** σ^2
Random Forest	0.005	0.99
OMP	0.04	0.97
Neural Network	0.35	0.38

β^***= 40 cm**

Model	**MAE** $[10^{-5}m^{-2}]$	**Explained** σ^2
Random Forest	0.005	0.99
OMP	0.21	0.76
Neural network	0.33	0.47

Comparison of preliminary results on different optics setting and data sets shows that Random Forest algorithm achieves the most accurate prediction. The presented prototype demonstrates that a machine learning based approach is promising and produces meaningful and consistent results. For further improvements on prediction quality, different model boosting techniques which construct ensembles with higher capacity than individual models have to be investigated. The simulation of training datasets including different machine settings is required to train the model appropriately and to ultimately produce predictions that could be used during optics commissioning and machine developments.

CONCLUSIONS

The suitability of machine learning methods has been clearly shown in the performed experiments. Considering future accelerator projects that will require more challenging optics control, more powerful analysis methods will be needed - machine learning techniques might become the key technology to implement fast powerful data analysis and discover new useful relations and features in the data. The positive results presented in this work open more space for further investigation and deeper analysis of appropriate approaches.

ISBN: 978-3-95450-184-7 doi:10.18429/JACoW-IPAC2018-WEPAF062

REFERENCES

[1] P. Baldi, P. Sadowski, and D. Whiteson, "Searching for exotic particles in high-energy physics with deep learning", Nature Commun. **5**, 4308 (2014).

[2] A. Aurisano *et al.*, "A convolutional neural network neutrino event classifier", JINST **11** 09, P09001 (2016).

[3] B. Everitt, S. Landau, M. Leese, and D. Stahl, " Cluster analysis", 5th ed. Chichester: Wiley (2011).

[4] A. Langner and R. Tomás, "Optics measurement algorithms and error analysis for the proton energy frontier", Phys. Rev. ST Accel. Beams **18** 031002 (2015).

[5] T. Bach and R. Tomás, "Improvements for optics measurement and corrections software", CERN-ACC-NOTE-2013-0010 (2013).

[6] R. Calaga and R. Tomás, "Statistical analysis of RHIC beam position monitors performance", Phys. Rev. ST Accel. Beams **7** 042801 (2004).

[7] L. Malina *et al.*, "Performance optimization of turn-by-turn beam position monitor data harmonic analysis", presented at IPAC'18, Vancouver, Canada, paper THPAF045.

[8] Stuart P. Lloyd, "Least squares quantization in PCM", IEEE Transactions on Information Theory **28**, p. 129-137 (1982).

[9] M. Ester, H. Kriegel, J. Sander, X. Xu, "A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, AAAI Press, 1996.

[10] R. Tomás, T. Bach, R. Calaga, A. Langner, Y. I. Levinsen, E. H. MacLean, T. H. B. Persson, P. K. Skowronski, M. Strzelczyk, G. Vanbavinckhove, and R. Miyamoto, "Record low β beating in the LHC", Phys. Rev. ST Accel. Beams **15** 091001 (2012).

[11] A. Wegscheider, A. Langner, R. Tomás and A. Franchi, "Analytical N beam position monitor method", Phys. Rev. Accel. Beams **20**, 111002 (2017).

[12] T. Persson, F. Carlier, J. Coello de Portugal, A. Garcia-Tabares Valdivieso, A. Langner, E.H. Maclean, L. Malina, P. Skowronski, B. Salvant, R. Tomas, and A.C. Garcia Bonilla, "LHC optics commissioning: A journey towards 1% optics control", Phys. Rev. Accel. Beams **20** 061002 (2017).

[13] R.Tomás, M. Aiba, A. Franchi, U. Iriso, "Review of linear optics measurement and correction for charged particle accelerators", Phys. Rev. Accel. Beams **20** 054801 (2017).

[14] E. Fol, "Evaluation of machine learning methods for LHC optics measurements and corrections software", Master thesis, CERN-THESIS-2017-336.

[15] L. Breiman, "Random forests", Mach. Learn. **45** (2001).

[16] R. Rubinstein, M. Zibulevsky, M. Elad, "Efficient Implementation of the K-SVD Algorithm using Batch Orthogonal Matching Pursuit", Cs Technion **40** (2008).

[17] C.M. Bishop, "Neural networks for pattern recognition", Oxford University Press, USA (1995).