Automated Machine Learning for Grid Diagnostics using Power Line Communications

Yinjia Huo, Gautham Prasad, Lutz Lampe and Victor C.M. Leung

Department of Electrical and Computer Engineering, The University of British Columbia

E-mail:{yortka, gauthamp, lampe, vleung}@ece.ubc.ca

Index Terms

Grid diagnostics, automated machine learning (autoML), cable aging

I. Extended Abstract

SINCE power line communications (PLC) use the existing grid infrastructure for data transmission, it is favorable to reuse the employed power line modems (PLMs) also for grid diagnostics [1]–[3]. As the cable ages, the dielectric properties of its insulation continuously deteriorate [4, Ch. 5]. This deterioration also manifests as a change in the PLC channel [2], which can be captured during the channel estimation procedure inside the PLMs. Therefore, the grid can be monitored remotely simply by extracting information indicative of cable health conditions from the estimated PLC channel.

However, a challenge that arises with this method is to distinguish PLC channel changes that are produced by cable degradations and the variations caused due to other activities, for e.g., varying load conditions. To this end, in our previous work, we proposed a data-driven framework for cable diagnostics, as shown in Fig. 1 [1]. We begin by identifying a salient localized degradation on each network branch. If no such degradation is present, we proceed to predict an equivalent cable age that provides an intuitive indication into the homogeneous degradation severity level of the cable. On the other hand, when we encounter a salient localized degradation, we assess its condition by predicting its extent as well as its degradation severity level. To achieve these goals, we use machine learning (ML) techniques to train either a classifier or a regressor for automated diagnosis.

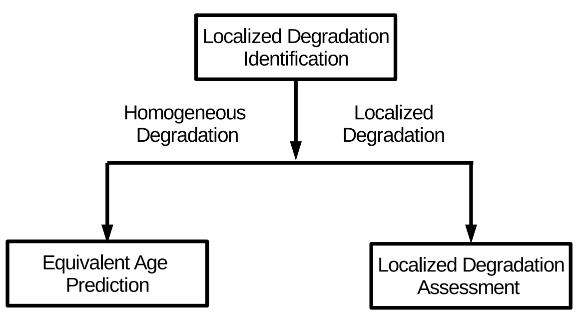
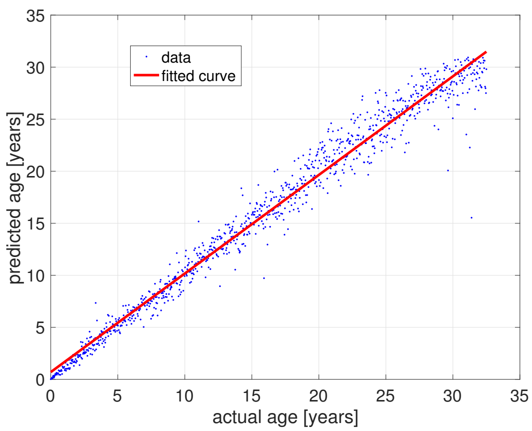


Fig. 1: Cable Diagnostics Framework

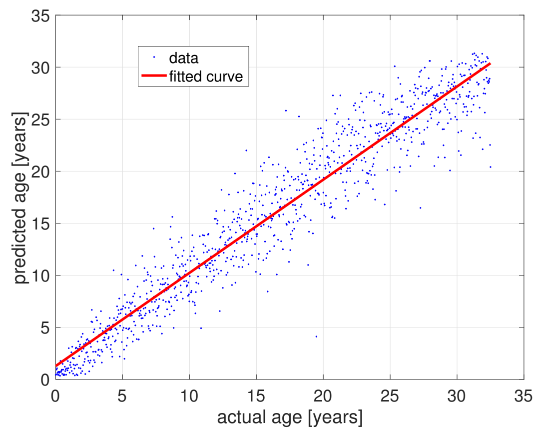
In this paper, we build on our previous work to investigate the use of the automated ML (AutoML) technique that automatically selects the best ML algorithm and the associated hyper-parameters to optimize performance in terms of the user-defined cost function for the given ML task within a fixed computational budget. While in [1] we suggested the use of boosting methods that provide excellent performance when each training/testing sample contains small number of features [5, Table 16.3], it does not guarantee the best performance for each of the tasks in our proposed cable diagnostics framework (see Fig. 1). At the same time, an investigation into the use of different ML algorithms could also clarify the abilities and limitations of our proposed cable diagnostics framework.

For the implementation of AutoML, we use auto-sklearn that enables meta-learning and automated ensemble construction of models for increased efficiency and robustness [6]. We apply auto-sklearn to each ML task of our cable diagnostics framework. In this work, we investigate the water-treeing degradation in extruded cross linked polyethylene cables.[1] For a fair comparison with the state-of-the-art, we use the same training and testing samples as those in [1].

[1]Please refer [1] for the water-treeing model, bottom-up PLC channel generation procedure, applied network topology and load conditions, and methods used to generate training and testing channel conditinos.



(a) trained with the raw data



(b) trained with the extracted features

Fig. 2: Equivalent age prediction for homogeneous cable degradations using auto-sklearn.

The results of equivalent age prediction for homogeneous cable degradations achieved by auto-sklearn is shown in Fig. 2. For the results of Fig. 2a, we use the raw channel impulse response (CIR) data to train the machine without any feature extraction. Next, we test the performance of AutoML by feeding in manually extracted features to determine its impact on the performance. Fig. 2b shows the results for this setting, where we use the same set of features extracted in [1] to train the machine.

We notice that for the task of equivalent age prediction, training the machine with raw data gives the best performance both in terms of prediction accuracy and variance. The fitted line nearly passes through the origin with a unit slope, and each individual prediction lies close to the fitted line. Specifically, we obtain a root-mean-square deviation (RMSD) of 1.52 by training with the raw CIR data (Fig. 2a), while the RMSD is as high as 2.69 when we manually feed in extracted features (Fig. 2b). For comparison with the state-of-the-art, least-square boosting used in [1] achieves an RMSD of 3.46. Therefore, AutoML, which uses a combination of different ML algorithms to construct an ensemble of models achieves close to 60% lower RMSD than using a single ML algorithm applied in [1]. In particular, the results of Fig. 2b are achieved with a combination of adaptive boosting, gradient boost, gradient boost (with different hyper-parameters), extra trees, decision trees, and k-nearest neighbors, where the proportion of each of them are 0.52, 0.22, 0.1, 0.08, 0.04, and 0.02, respectively.

While we significantly improve the prediction performance using the constructed ensemble of different models, we recognize that our proposed solution introduces higher computational cost. Further, for some tasks, for e.g., predicting the length of a localized degradation (not shown in this extended abstract), we notice that training the machine with extracted features (as done for the results of Fig. 2b) provides better prediction accuracy compared to training with raw data. This shows that domain knowledge can sometimes assist us in selecting useful features for a specific ML task.

In conclusion, along with generally satisfactory performance, AutoML has also helped us gain deeper understanding into the suitability of different ML algorithms for each of the cable diagnostics tasks.

Acknowledgment

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] Y. Huo, G. Prasad, L. Atanackovic, L. Lampe, and V. C. M. Leung, "Grid surveillance and diagnostics using power line communications," in *Proc. IEEE Int. Symp. Power Line Commun. Appl. (ISPLC)*, 2018, pp. 1–6.

[2] L. Förstel and L. Lampe, "Grid diagnostics: Monitoring cable aging using power line transmission," in *Proc. IEEE Int. Symp. Power Line Commun. Appl. (ISPLC)*, 2017, pp. 1–6.

[3] F. Passerini and A. M. Tonello, "Full duplex power line communication modems for network sensing," in *IEEE Int. Conf. on Smart Grid Commun. (SmartGridComm))*, 2017, pp. 1 – 5.

[4] G. Mugala, "High frequency characteristics of medium voltage xlpe power cables," Ph.D. dissertation, KTH, 2005.

[5] K. Murphy, *Machine Learning: A Probabilistic Perspective*, ser. Adaptive computation and machine learning. MIT Press, 2012. [Online]. Available: https://books.google.ca/books?id=NZP6AQAAQBAJ

[6] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," in *Advances Neural Inform. Process. Syst. 28*. Curran Associates, Inc., 2015, pp. 2962–2970. [Online]. Available: http://papers.nips.cc/paper/5872-efficient-and-robust-automated-machine-learning.pdf