21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2021 Grid of the Future Symposium

Practical Machine Learning Applications

T. RHODES[1], I. MITICHE[2], T. MCGRAIL[3], F. WERNER[4], P. BOREHAM[5]
Duke Energy[1], Glasgow Caledonian University[2], Doble Engineering[3,4,5]
USA[1,3,4], United Kingdom[2,5]

SUMMARY

The role of analytic tools to provide diagnostic insight into substation asset viability is becoming crucial: resulting actions may be taken autonomously, as happens at present with protection schemes, but the reasons for action or intervention must be both auditable and justifiable as they may have serious consequences. This paper looks at some of the practical issues with implementing machine learning (ML) approaches to data analytic development, from slowly varying dissolved gas analysis (DGA) data through to voluminous phase resolved partial discharge (PRPD) data. An advantage of ML is the ability to use large data sets as a training reference from which the virtual machine learns. However, there is a need for the machine to maintain 'focus', that is, for the machine to be trained to perform a specific task and not expect it to be as capable or accurate as data becomes less 'familiar'. Interpolating between members of a large data set to 'identify' new exemplars is far more likely to succeed than in extrapolating from the original data. Experiences with DGA results and related data to provide inputs to health scoring systems described here have both pro's and con's. In addition, the analysis of PRPD data is used to show the benefit of checking whether a test case is 'close enough' to the training data before making an interpretation, and thus avoid one route to mis-classification.

KEYWORDS

Power Transformer, Machine Learning, Artificial Intelligence

tmcgrail@doble.com

INTRODUCTION

An advantage of machine learning (ML) is the ability to use large data sets as a training reference from which the machine learns. However, there is a need for the machine to maintain ‘focus’, that is, for the machine to be trained to perform a specific task and not expect it to be as capable or accurate as data becomes less ‘familiar’. The ability of ML to address interpolation may be far greater than the ability to address extrapolated data, or outliers.

We use the background of transformer dissolved gas analyses to show that the interpretation of data is not straightforward. Dissolved gases in transformer oil are measured and used for both detection and diagnostic purposes:

- Detection of possible faults within the transformer through rising gas levels
- Diagnosis of the nature of the fault through analyses of gas ratios and proportions of each gas within the overall sample.

Rising hydrogen levels, when viewed in isolation, may indicate a number of possible faults, such as PD, or overheating: rising hydrogen level is thus a useful fault detection ‘analytic’. If hydrogen is rising in conjunction with acetylene, that could indicate something far more serious: possibly an arc with power follow through could be diagnosed. There are a number of different diagnostic schemes available, some dating back several decades, for the interpretation of the dissolved gas levels: from Rogers’ Ratios and Dornenburg Ratios, through the triangles and pentagons of Michel Duval, to CIGRE or IEC interpretation guides [1]. One downside of the availability of several interpretation schemes is that they do not always indicate the same diagnosis, and ‘final’ interpretation of data may require expert analysis. There may also be variation in interpretation based on the manufacturer/design/vintage of each transformer, all of which may need to be taken into consideration. Over several decades, ML and artificial intelligence (AI) approaches have often been applied to dissolved gas analysis data, with varying degrees of success [2].

DUKE ENERGY HYBRID ML APPROACH

Duke energy has applied machine learning in a targeted and systematic way to give confidence in the results and to show benefits of ML application to others within the asset management organization. An initial step was to compare the performance of an ML system in identifying improvements in transformer condition based on work orders (WO) for maintenance activities performed: historical data was used to generate health indices, and those were shown to improve after maintenance of the cooling system of a large population of transformers. This improvement could, of course, be misleading if the improvement seen by the ML did not relate to the actual condition of the units, for example being solely related to the WO actually being performed at some point: condition data is needed to show the benefit, in this case from application of DGA standards/guidelines. In addition, a ‘live test’ was performed using a subject matter expert (SME) led ‘manual’ analysis of 650 open work orders on power transformers, again including application of DGA standards and guidelines to identify those units which may have a Duval interpretation of severe cooling issues, of a ‘T1’ or ‘T2’ classification, and, in parallel, the ML approach was also applied to review the same data. There was agreement between the two sets of results in that the SME manual analysis and the ML analysis both in identified 36 units of particular interest but, further, the ML approach flagged up a number of borderline cases which could otherwise have deteriorated and been missed; application of simple rules and standards/guidelines would not have identified the borderline units.

In the work at Duke Energy, the benefit of applying ML is one where it supports the SME's and allows them to concentrate on units of interest, rather than trawling through vast amounts of data to identify targets. The ML system has a large data set on which to focus, but the SME involvement ensures the learning is closed loop, and the results give confidence in the application of ML. Each asset is classified as being in one of several states: 'stable' meaning that the condition should not deteriorate in the immediate future, 'service' , as the name suggests, meaning that a practical intervention should be scheduled, and so on. When the 'state' of an asset is moved from 'stable' to 'service' or 'risk identified' an SME is required to perform the resetting after appropriate intervention; the ML does not get to say everything is 'back to normal'. The result is fewer things 'slipping through the cracks' as each SME has over a thousand power transformers to manage on a daily basis, along with associated breakers and equipment.

It should be noted that SMEs at Duke believe that 95% of the benefit of ML can be achieved through application of data clean up, ensuring data consistency, and systematic application of rules/standards/guidelines, but that the extra 5% is based on borderline cases being identified early enough to support intervention in a meaningful manner. The ML is also focused – required to do well defined tasks on well organized data, and has gained acceptance widely within the organization.

Duke energy have developed ML approaches which apply standard engineering analysis tools for dissolved gases, but extend those to add in more 'context' data in an ML environment – the result is a system which is far more effective at identifying 'benign' results and allowing the experts to concentrate on 'questionable' results which may require resampling or further testing to identify an underlying problem. As the number and variety of data sets included in the ML system have grown, Duke Energy have shown that they can benefit from this 'hybrid' approach, focusing valuable and scarce resources on those transformers which are not only most in need of attention, but also of most risk to the system.

PHASE RESOLVED PD INTERPRETATION

Analysis of PD data can be a significant challenge with interpretations often requiring a combination of experience and suitable data presentation for the expert to be able to provide an interpretation []. Classifying phase resolved PD patterns using ML can provide a means to support interpretations in the field and focus on those which seem to provide an indication of a serious problem. The approach discussed here is a complex analysis which allows for identification of the similarity of the pattern of interest to the original training data – thus giving confidence that the subsequent classification is 'valid'. An explanation is provided for both the 'out-of-distribution' (OOD) calculation and the subsequent pattern identification.

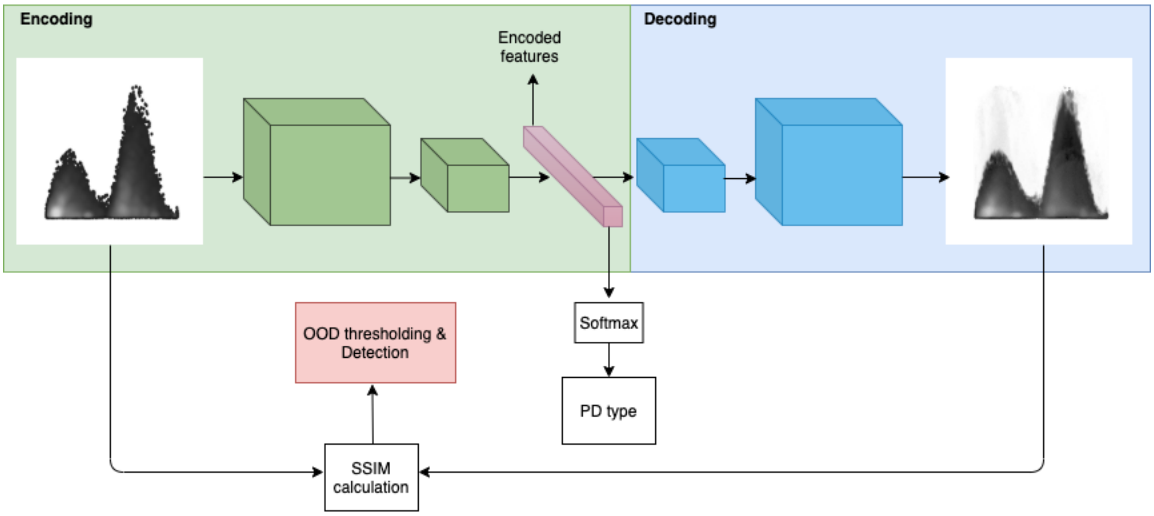

Figure 1 Deep learning framework for PRPD classification and out-of-distribution detection

The ML system for PD classification in rotating machines is able to identify and classify EMI/PD types in rotating machines using PRPD patterns as input to a deep learning model. In addition, an of OOD detection approach is implemented to prevent false classification of data that is not similar to the EMI/PD types used in training the deep learning model. This process is summarised in figure 1, where the PRPD image is obtained through density scatter plot of PRPD data. The images from 7 EMI/PD types in rotating machines are used to train the model for classification and image reconstruction in an auto-encoder approach. The Structural Similarity Index Measure (SSIM) is calculated between the original and reconstructed images across the training data and thresholded for OOD decision.

The PD type classification is obtained using a Softmax function on the encoded features defined as:

$$S(\boldsymbol{x}) = \frac{e^{\boldsymbol{x}}}{\sum_j e^{\boldsymbol{x}_j}}$$

This function highlights the large values in the vector x and minimises the smaller values, resulting in rescaled values to a range of (0,1) which sum to 1.

The OOD decision is based on comparing the SSIM value to a threshold derived from the distribution of the training data SSIM values. The SSIM is a reconstruction metric that measures how close the reconstructed image $\tilde{X}$ is to the original image X. It can be calculated as:

$$SSIM = \frac{(2\mu_X\mu_{\tilde{X}} + c_1)(2\sigma_{X\tilde{X}} + c_2)}{(\mu_{\tilde{X}}^2 + \mu_{\tilde{X}}^2 + c_1)(\sigma_{\tilde{X}}^2 + \sigma_{\tilde{X}}^2 + c_2)}$$

Where μ and σ are the mean and standard deviation respectively, c_1= (0.01L)^2 and c_2=(0.03L)^2 where L=$2^{(\#bits/pixel)}$-1 is the range of the 255 pixel values for 8-bit grayscale image. A perfect match between two images results in a value of 1 and reduces toward 0 as the reconstruction quality degrades.

The proposed deep learning model was tested with real-world Electromagnetic PD data measured at a 106 MVA General Electric (GE) steam turbine generator, specifically at the stator winding, using the 2nF couplers installed on phases A, B, C as well as the neutral connection of the generator. Figure 2 shows example PRPD data (left) and their respective classification result (right) by the Softmax of the deep learning model. The SSIM metric for (a) is 0.94 and for (b) is 0.38, since the derived threshold is 0.6 (a) is accepted as in-distribution and (b) as OOD. These results have been confirmed to be correct by a PD expert. For (b) it is observed that the deep learning model's output identifies which features from the 7 PD types are present in the test data (c) which is surface discharge and end-winding PD, however according to the PD expert the data contains multiple PD defects. This is when the OOD detection provides a level of confidence in the classification.

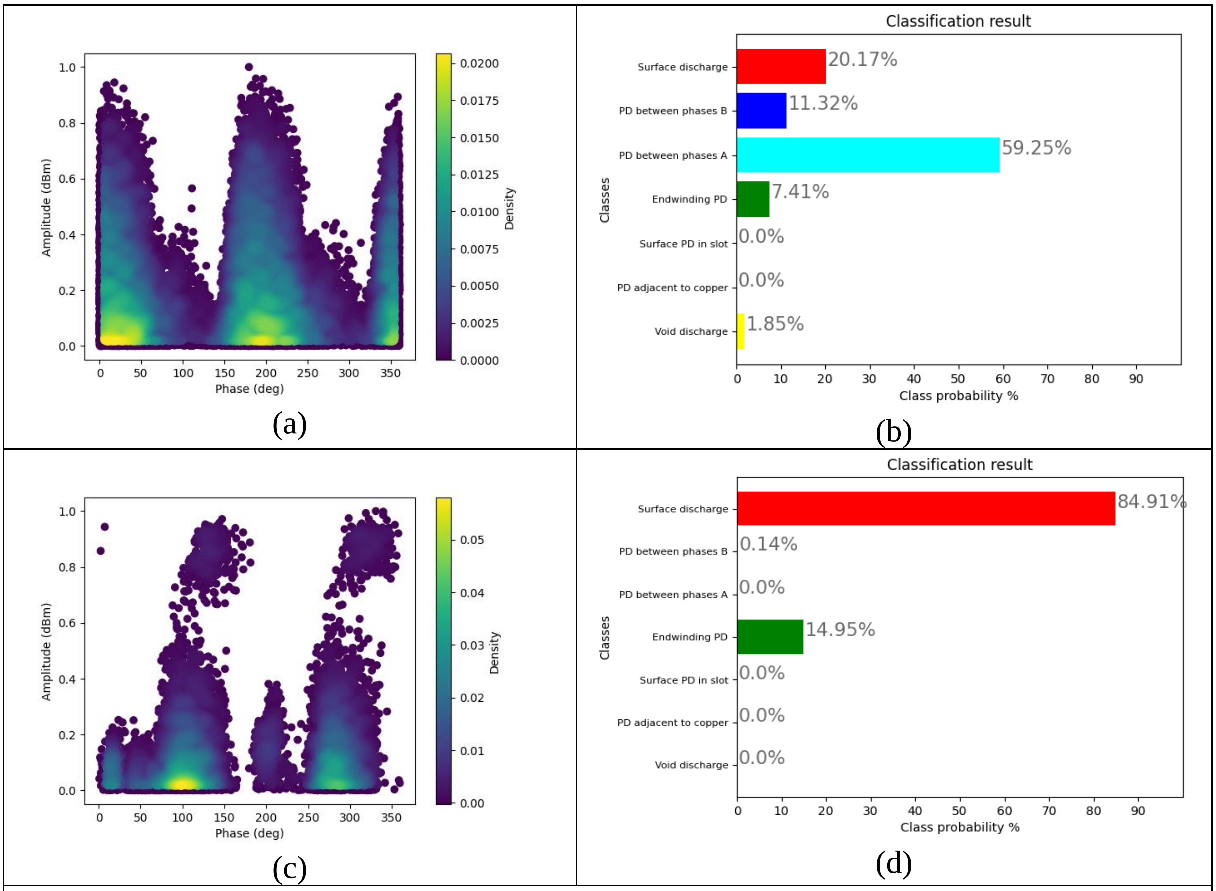

Figure 2. Example PRPD data used in testing the model and classification results.

The successful application of both ML and deep-ML is predicated on having a suitably representative data set, otherwise mis-classifications can result. The deep-ML application for PRPD is extended by first checking the 'similarity' of a new case to the original data set and thus reducing the possibility of mis-classification. Subsequent classification includes the possibility of a 'new class' which supports the development of larger and more representative data sets.

DISCUSSION

The application of ML systems is becoming more common in the electric supply industry, and their value is often in performing well defined tasks repeatably and reliably. There is still a role for expertise at this point, where data for interpretation does not well match the original training data for the ML system. The size of that role may well shrink over time as the available data for analysis grows.

In identifying DGA results of 'interest' Duke Energy has supported practical field engineers and colleagues by helping them focus on those 'interesting' results. The hybrid approach does not replace engineers but helps them focus where they bring most value

In PRPD the classification of patterns improved when moving from ML to deep-ML, but also benefits from the OOD, which helps identify those cases which may not be represented in the original data set: the likelihood of mis-classification can thus be reduced, and confidence in the deep-ML application increased.

ACKNOWLEDGMENTS

The authors wish to thank their many colleagues for free and frank discussions on the use and abuse of machine learning systems which have helped focus and improve this work.

BIBLIOGRAPHY

[1] N. A. Bakar *et al*, "A Review of Dissolved Gas Analysis Measurement and Interpretation Techniques" (IEEE DEIS Magazine, Vol 30, No 3, pages 39-49)

[2] Y. Benmahamad *et al*, "Application of SVM and KNN to Duval Pentagon 1 for transformer oil diagnosis" (IEEE Transactions on Dielectrics and Electrical Insulation, Vol 24, No 6)