NIA ENWL020

Machine Learning

Closedown Report

A Network Innovation Allowance Project

31 July 2022

VERSION HISTORY

Version	Date	Author	Status	Comments
V0.1	22.07.22	Ben Ingham	Final	

REVIEW

Name	Role	Date
Lucy Eyquem	Innovation PMO Manager	28.07.22
Ben Ingham	Innovation Technical Manager	22.07.22
Victoria Turnham	Head of Innovation	25.07.22

APPROVAL

Name	Role	Date
Steve Cox	DSO Director	28.07.22

CONTENTS

1 EXECUTIVE SUMMARY 5

2 PROJECT FUNDAMENTALS 5

3 PROJECT BACKGROUND 5

4 PROJECT SCOPE 5

5 OBJECTIVES 6

6 SUCCESS CRITERIA 6

7 PERFORMANCE COMPARED TO THE ORIGINAL PROJECT AIMS, OBJECTIVES AND SUCCESS CRITERIA 6

8 THE OUTCOME OF THE PROJECT 6

9 REQUIRED MODIFICATIONS TO THE PLANNED APPROACH DURING THE COURSE OF THE PROJECT 9

10 PROJECT COSTS 9

11 LESSONS LEARNT FOR FUTURE PROJECTS 10

12 PLANNED IMPLEMENTATION, RECOMMENDATIONS OR NEXT STEPS 10

13 DATA ACCESS & QUALITY DETAILS 10

14 FOREGROUND IPR 10

15 FACILITATE REPLICATION 10

16 STANDARDS DOCUMENTS 10

17 OTHER COMMENTS 10

GLOSSARY

Term	Definition
ML	Machine learning
AI	Artificial intelligence
DSP	Digital Signal Processing
LV	Low voltage
LCT	Low carbon technologies

1 EXECUTIVE SUMMARY

The aim of this project was to collate the various data steams being collected by ENWL form installed low voltage network and transformer monitoring devices and to explore whether the application of Machine Learning and Artificial Intelligence could lead to enhanced learnings.

The project examined the existing data and also developed new triggering mechanisms to capture additional information. New algorithms were then developed to process this data and understand what additional value could be extracted. Over the course of the project it was shown that by applying these techniques noticeable improvements could be made to the fault location algorithms. The additional data gathered from the LV monitoring equipment, along with the new triggering mechanisms, provided more detail on the network conditions and could be used to detect the presence of LCTs. Finally the revised processing of the transformer monitoring data is able to detect additional signs of developing issues

2 PROJECT FUNDAMENTALS

Title	
Project reference	NIA_ENWL020
Funding licensee(s)	Electricity North West Limited
Project start date	October 2018
Project duration	3 years
Nominated project contact(s)	innovationteam@enwl.co.uk

3 PROJECT BACKGROUND

Finding an alternative way to interrogate this data could be crucial to the future management of DNO Large volumes of data have become available over the last few years due to the intelligent devices fitted to the network as BAU or innovation projects. This large dataset holds information regarding to network operation and performance, asset health, development of faults and abnormalities on the network. Analysis of this data is currently a time-consuming manual process so only clearly defined small pieces of analysis takes place. The data may hold hidden trends which could not be manually investigated and may offer valuable insight into network operations which could influence investment decisions as well as response to events.

networks. This project proposes to investigate whether modern techniques such as machine learning and artificial intelligence could assist with this interrogation.

4 PROJECT SCOPE

This project will be a research piece investigating the application of machine learning and artificial intelligence to data already being collected by low voltage monitoring equipment and transformer monitoring equipment already deployed on the network. The research will investigate whether machine learning can be used to identify hidden trends and make recommendations for network investment.

5 OBJECTIVES

This project will look to collate data from the various systems currently in use across Electricity North West. Build, train and evaluate a model to classify and work with the data. Use this model to produce recommendations for network operation and investment.

6 SUCCESS CRITERIA

- Production of a report on the methodology for collating the data sets
- Production of a model to interrogate the data sets
- Report detailing outputs from the model and recommendations for network operation and investment
- Report detailing how the model can be transferred to business as usual

7 PERFORMANCE COMPARED TO THE ORIGINAL PROJECT AIMS, OBJECTIVES AND SUCCESS CRITERIA

The project partner resources required for the delivery of the project were identified and put in place, with many of the resources being supplied by the Camlin Machine Learning (ML) and Artificial Intelligence (AI) research centre, based in Parma, Italy. This group of researchers bring vast experience of applying ML/AI techniques to real world datasets, having been previously recognised by MIT for their work.#

Additional resource from the Camlin Digital Signal Processing (DSP) and Firmware team was put in place to enable new firmware to be developed, tested and deployed to thousands of network deployed low voltage (LV) intelligent electronic devices (IEDs) to collect additional data.

7.1 Additional Low Voltage (LV) Data Collection Activities

During the project data collation activities gathered large volumes of information from low voltage networks, as well as tap changer monitoring and transformer monitoring, and was transferred into a streamlined system implemented and utilised by the Parma group.

In addition, new triggering methods for field deployed equipment have been developed by the DSP and Firmware engineering team, allowing additional information outside of the standard triggering/data gathering scope to be captured for analysis. This has undergone rigorous testing before being deployed to many field devices. The new data collected by the new triggering methods will improve understanding of the signatures from the loads on the network.

Low Voltage devices in the network such as BIDOYNG and WEEZAP are currently producing a large dataset in respect to LV faults and perturbations, as well as 30-minute power quality data. This is very valuable data but is missing one key component valuable for machine learning data analysis methods. There are no high bandwidth records from a period on the network when there are no perturbations. The waveforms currently being recorded do contain a few cycles of 'pre-fault' data, but this may contain information relating to the onset of the perturbation that is also recorded with the record. This work package developed new triggering methods that will gather high bandwidth data captured on the network during normal operation. Existing triggering methods are based on perturbations such as over current or under voltage methods. Time or random based data capture triggering will allow a model of normal behaviour to be built up that will enable new innovative investigation methods on the existing monitoring devices on the network.

Data captured during the normal operation of the network will help with understanding the signatures from network loads and will improve the interpretation of data from spike and step-based triggers used in sensitive applications, such as Rising and Lateral Mains.

Data was collected for a full year using the updated triggers developed as part of this project and deployed as a firmware update to several devices installed in the network. This data collection was and continues to be essential to the modelling and analysis activities of this and other projects.

7.2 LV Model Development

The aim of this work stream is to take the data collected relating to normal behaviour from the first work stream, and to inform and construct models of the low voltage network for a several tasks with valuable outcomes.

The need for three specific types of LV network models was identified. The first model is being developed and tuned to understand the classification of different types of loads on individual feeders, as well as predict growth of the load over time.

A second model uses information from standard monitoring devices to perform a basic health assessment and ranking of cables, and a third will be developed to identify among the data where and when Low Carbon Technologies are utilised.

7.3 Categorisation and classification of load profiles

An advanced LV model for classification of load types attached to each feeder has been created and run on a large data set, producing a total of 16 individual classifications, with 12 uniquely identifiable and labelled with their type, and 4 highly mixed load types.

The second model uses device information to perform a basic health assessment and ranking of the health of cables, and a report including the ranking of networks is now available through Tableau for Rising and Lateral Mains installations, helping to prioritise investment and maintenance resources.

The third model was developed by a team specialising in machine learning, targeted firstly at identifying existing issues created by low carbon technologies. Continuation of this work will be used to predict feeders and network areas that will likely develop problems in the future, such as capacity or voltage issues.

Network modelling will enable results for one region to be applied to others, identification of sections of network that are compatible (both physically and based on load patterns) for integration, meshing and automation. Identification of growth of specific types of devices connected to the network, including EV, heating, PV etc. Prediction of patterns in the models – allowing the management of significant penetration of EV's and heat, provide forecasting methods for changing behavioural patterns, and derive better asset performance and lifetime assessment. This will also allow the management of tighter capacity margins. This work will also likely lead to identification of hitherto unknown patterns and relationships in the data, which may lead to new insights into network operation.

7.4 Fault Management

Improvements to fault management systems are ongoing, with fault classification and labelling work dependent on feedback and collation of data about actual fault locations. Kelvatek have previously processed data using a Variational Autoencoder, a type of Deep Neural Network, to experiment with moving from a higher dimensional representation of LV fault data to a lowdimensional simpler structure containing the same information.

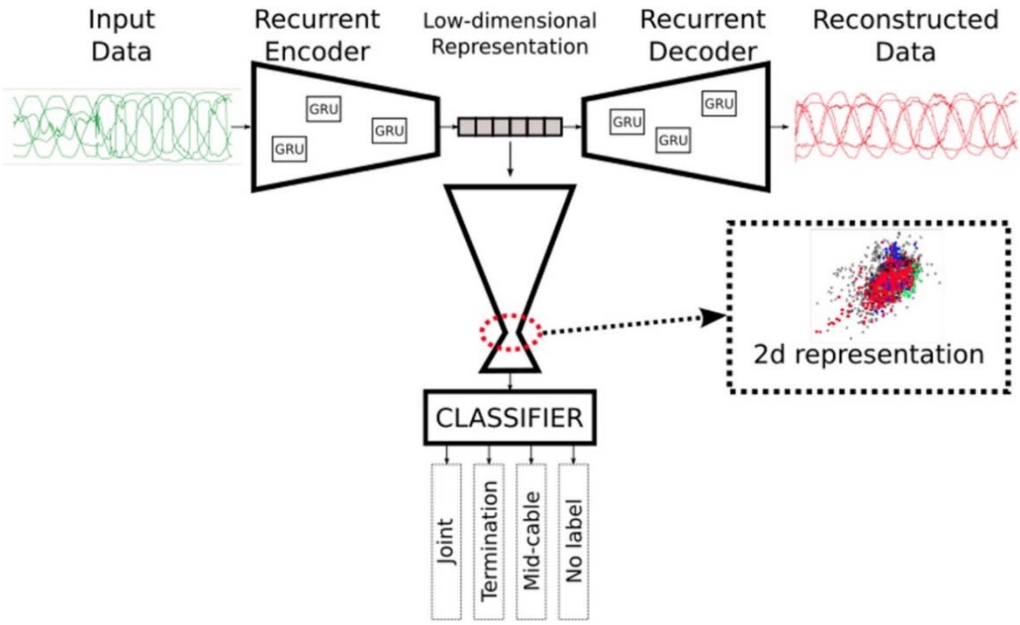

Figure 1 - Variational Autoencoder fault classifier

The outcome has shown that a machine learning based approach can help distinguish between mid-cable faults and faults that are either termination or joints. The percentage of successful classification between joints & terminations against mid-cable is 92%. As the database grows the accuracy of the determination is expected to help improve the identification of joint faults from those located in terminations.

This work package concentrated on the development in this algorithm, and the collection of additional data, for selecting the location of faults from between mid-cable, terminations and joints in cable networks from fault records. It will also strive to provide a practical way to apply this analysis to each 'Go Locate' job raised by Kelvatek's SAPEINT centre to localise LV cable faults.

7.5 Tap Changer Monitor Data Analytics

It is estimated that there is around 5000 onload tap changers (OLTC) in the UK distribution network. Many different designs exist, and there are several variations within the internal mechanisms. They allow a momentary diversion of the load being carried by a transformer to allow a change to be made to the number of turns in a transformer winding, there by changing the output voltage. OLTC's, like may electromechanical devices with stored energy mechanisms, are subjected to mechanical stresses that can lead to stress and fatigue fractures within the mechanism that cannot be easily detected during routine maintenance. These fractures can lead to catastrophic failure. Estimates show that up to 5 OLTC failures across the UK each year, and at least one of these failures will lead to the loss of the transformer as well as the OLTC. This work package seeks to take the data and algorithms created in the separate project called "On-Load Tap Changer Monitoring" and advance it to provide automated, real-time analysis of the data to alarm anomalies soon after they occur within the tap changer.

A separate project investigating the monitoring of tap changer operations by collating vibrational and electrical information during the operation of tap changers has provided data from over 170,000 events, with more than 1.5M waveforms collected. A machine learning model has been developed to analyse this data and build an anomaly detection system. A detection model was trained for each tap changer using its historical data and the model produces an anomaly score for each switchover event.

This model makes the identification of risky conditions and actual breakdowns possible, as well as spotting anomalous events and patterns not easily detectable during routine maintenance.

Whilst collecting the data from the tap changers, one unit developed a fault. Analysis was then carried out on the collected data which demonstrated that the faulted condition could have been detected prior to the fault occurring.

The work has shown that vibro-acoustic data can reveal relevant anomalies that cannot be seen with any other means.

7.6 Transformer Analytics

Electricity North West's innovative transformer oil regeneration programme aims to maximize the life span of transformers that are passing their original design lifespan. Through use of oil regeneration techniques combined with online monitoring and diagnostic systems an optimum regeneration programme can be developed. Kelvatek have participated in the project by supplying, installing and commissioning sophisticated monitoring systems (TOTUS), gathering the resulting data, and assisting ENW in the analysis of data to establish the optimum regeneration pattern to maximally extend transformer life. This project has been very successful so far, but it is believed that further insights will be possible through the use of machine learning methods to revaluate the data.

With the increase in the numbers of transformers that have online dissolved gas analysers embedded in the distribution network, a key question is whether we can use AI and ML methods to extrapolate asset management data and valuable key insights into their use and life spans.

8 THE OUTCOME OF THE PROJECT

As reported above there the project successfully demonstrated the value of the machine learning approach in several areas. In terms of the LV monitoring data the new triggering mechanisms allowed for additional information to be captured which can give a greater understanding on network conditions.

Additional work on the data collected from these devices was used to generate new, more accurate fault location algorithms. These will be further improved as additional, real world data is fed in as part of BaU usage.

9 REQUIRED MODIFICATIONS TO THE PLANNED APPROACH DURING THE COURSE OF THE PROJECT

There were no changes to the project methodology required.

10 PROJECT COSTS

Item	Category	Estimated costs (£k)	Final costs £k (rounded)	Variance

1	Project Management	£225k	£215k	£10k
2	Research & development	£600k	£612k	-£12k
	Total	**£825k**	**£827k**	**£-2k**

The project was closed down with an overall overspend of £2k (0.2%)

11 LESSONS LEARNT FOR FUTURE PROJECTS

The key learning point for this and future data projects was around the varied nature of the data sources and the need to establish a single point of truth in order to manage any conflicting records.

12 PLANNED IMPLEMENTATION, RECOMMENDATIONS OR NEXT STEPS

The fault location algorithms have been incorporated into the existing systems and will be further refined using the data for any incidents that occur. This is expected to further increase their accuracy. We will look to work with our asset management team to bring the learnings from the condition monitoring algorithms.

13 DATA ACCESS & QUALITY DETAILS

Electricity North West's innovation data sharing policy can be found on our website.

There is currently no data available from the project.

14 FOREGROUND IPR

The default IPR position has been applied to this project, however the is no foreground IPR associated.

15 FACILITATE REPLICATION

The outcomes of the project should be easily replicable. The data gathering equipment is commercially available and in common use across DNO's. The application of the algorithms may depend on some training data for DNO specific scenarios but in general should be universal

16 STANDARDS DOCUMENTS

N.A.

17 OTHER COMMENTS

None