University of Strathclyde Engineering

Rachel Stephen Mollel, University of Strathclyde, rachel.mollel@strath.ac.uk
Lina Stankovic, University of Strathclyde, lina.stankovic@strath.ac.uk
Vladimir Stankovic, University of Strathclyde, vladimir.stankovic@strath.ac.uk

Supergen Energy Networks

Early Career Researcher Workshop

Using Explainability Tools to Inform Non-Intrusive Load Monitoring Algorithm Performance: A Decision Tree Approach

1.Introduction

- Non-Intrusive Load Monitoring (NILM) is a method of estimating electrical energy consumption and operation condition of individual appliances by observing the total aggregated power measurements from the smart meter with granularity of 10s to 30 mins

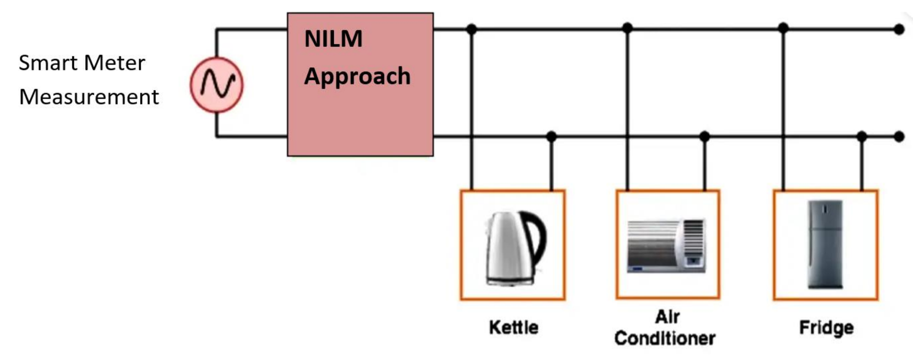


Figure 1: General description of non-intrusive load monitoring technique

- One of the benefits of load disaggregation is to help consumers apply energy-saving behavior to reduce living expenses and their carbon footprint
- This appliance information is likely to create demand for control systems, smart appliances, and demand response programs. In addition, appliance innovations with energy efficiency will be accelerated

3.Decision Tree Multi-Classifier

- Decision tree (DT) is a low-complexity supervised approach that requires only a small dataset to train the model. It has shown good performance for NILM and can be used effectively as a multi-classifier
- DT method is interpretable by design, in the sense that it is possible to design a tree in a way that decision outcomes can be mathematically explained and predicted

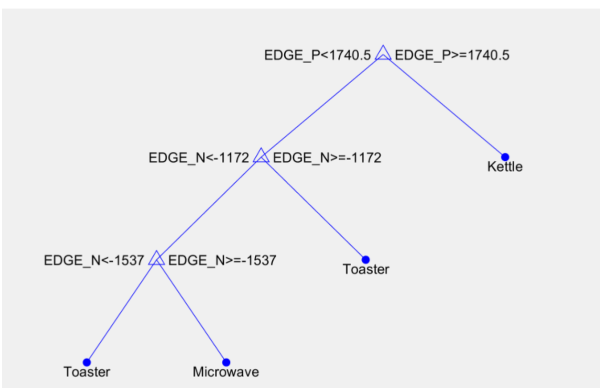


Figure 3: An example of a Decision Tree

- However, as the tree is becoming more complex with all the decision splits, the dependence of a predicted outcome on the feature is not easily seen
- Therefore, additional explainability methods are needed to shed light on most important features that steer the model towards certain decision

2.Explainability

- NILM model is said to be interpretable if the reason behind its prediction can be easily explained

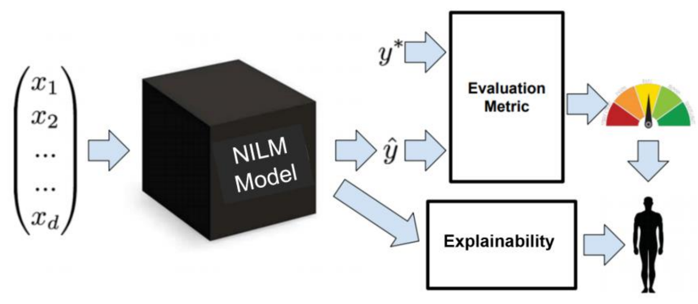
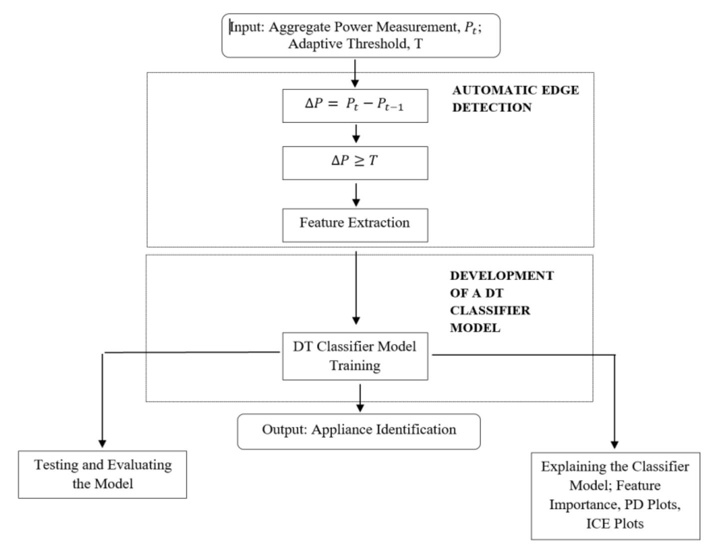


Figure 2: Explainability of NILM Model

- The importance of explainability of a NILM model is:
  - To facilitate learning and satisfy curiosity as to why certain decisions have been made by the model to build end user trust in AI NILM algorithm
  - For tuning purposes, as with explainability methods, one can learn important features that contribute significantly to the outcome and which do not, and
  - To debug the model in case of errors

4.Methodology



- The automatic event detection will output: (1) *EDGE_P*: ΔP value when the appliance became ON (in Watts), (2) *EDGE_N*: ΔP value when the appliance went OFF (in Watts) and (3) *DURATION*: time difference (in seconds) between time at *EDGE_P* and time at *EDGE_N*
- DT requires labelled data during training: The generated features will be used as input features with output labels (Appliances: Washing Machine, Dishwasher, Microwave, Toaster and Kettle) during training. Training is done on 55 edge-pairs per appliance taken randomly from the dataset (except during the testing months)
- After training, the model is exported for prediction on unseen data without labels. Testing is done on the entire unseen months of October, November, and December 2014 of the REFIT dataset
- For performance evaluation, the following standard classification metrics are used: Precision (PR), Recall (RE) and F-Score

5.Explainability Results

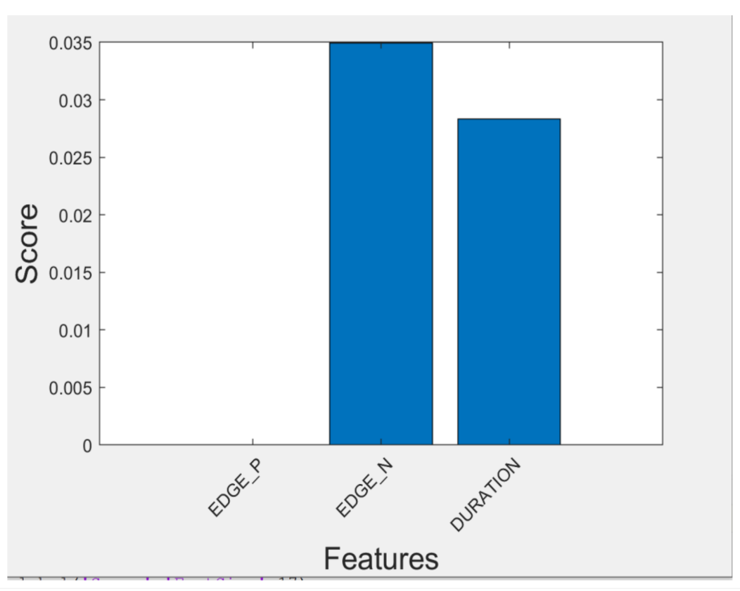


Figure 4: Feature Importance

APPLIANCE	PR	RE	F-SCORE
Dishwasher	0.84	0.61	0.71
Washing Machine	0.51	0.77	0.61
Kettle	1	1	1
Microwave	0.98	0.89	0.93
Toaster	0.61	0.90	0.73

	EDGE_N & DURATION			EDGE_N		
APPLIANCE	PR	RE	F-SCORE	PR	RE	F-SCORE
Dishwasher	**0.87**	**0.81**	**0.84**	0.71	0.60	0.65
Washing Machine	**0.69**	0.74	**0.72**	0.42	0.51	0.46
Kettle	0.98	0.99	0.98	0.94	0.95	0.94
Microwave	0.96	0.87	0.91	0.96	0.85	0.90
Toaster	0.61	0.90	0.73	0.57	0.90	0.70

Figure 5: Performance with Explainability-Informed Feature Selection

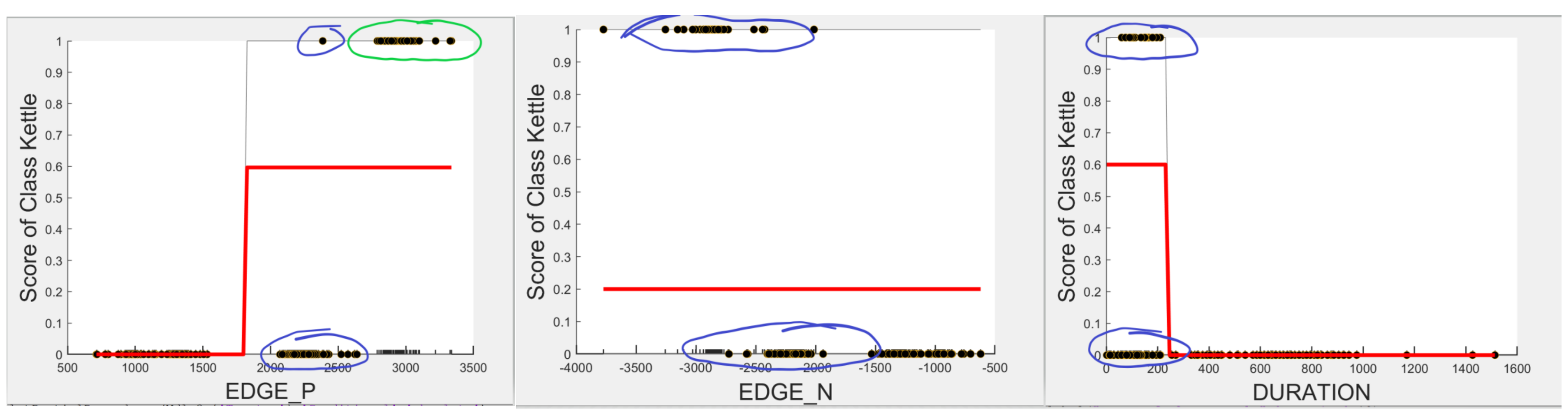


Figure 6: Performance with Explainability-Informed Feature Selection

6. Conclusion and Future Work

- This research proposes how explainability of a model yields a deeper understanding of the relative importance of features overall and on each instance of a prediction
- This in turn can be used to improve the model performance, in addition to improving the trustworthiness of the model
- We also show that explainability-informed feature selection improves performance of the classifier in general
- Exploring viability of transfer of models built on UK/EU/US datasets to the global south in terms of appliance level consumption feedback, demand forecasting especially in relation to presumption to reduce dependence on the grid

REFERENCE

- Rachel Stephen Mollel, Lina Stankovic, and Vladimir Stankovic. 2022. Using Explainability Tools to Inform NILM Algorithm Performance: A Decision Tree Approach. In 6th International Workshop on Non-Intrusive Load Monitoring (NILM '22), November 9–10, 2022, Boston, MA, USA. ACM, New York, NY, USA, 5 pages
- David Murray, Lina Stankovic, and Vladimir Stankovic. 2017. An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study. Scientific data. Scientific Data 4, 1 (01 2017). https://doi.org/10.1038/sdata.2016.122

https://www.ncl.ac.uk/supergenenhub