2022 CIGRE Canada Conference & Expo
Paper A2 - 505

Localization of Partial Discharge in Power Transformer Winding Using Sparse Autoencoder

H. Moradi Tavasani, S. Mantach, M. Gunawardana, B. Tabei, and B. Kordi

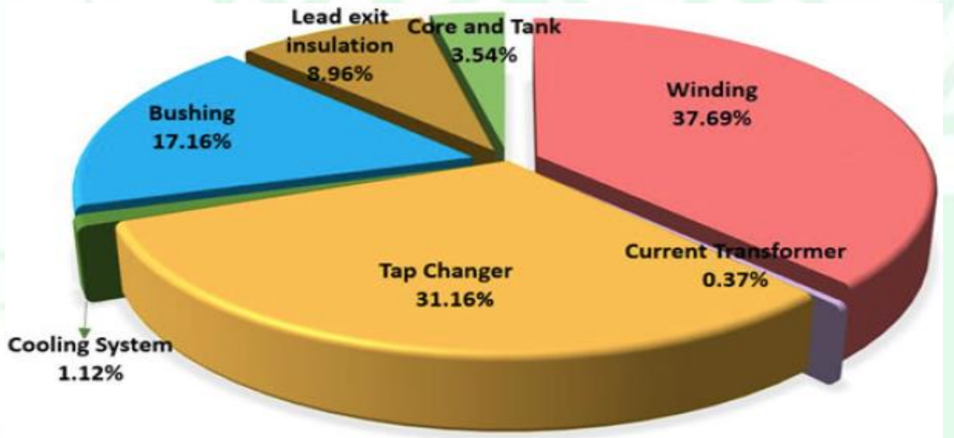
Department of Electrical and Computer Engineering

University of Manitoba

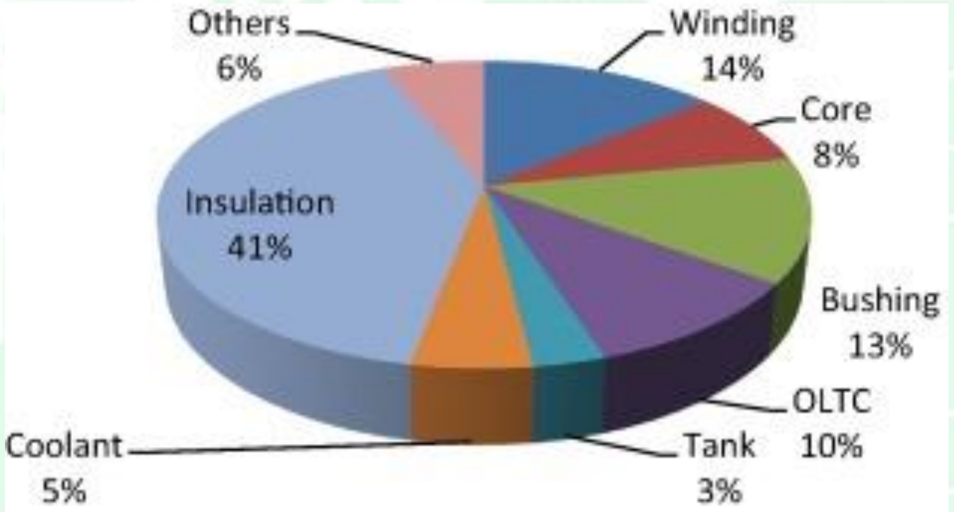
Winnipeg, Manitoba, Canada

Power Transformers Failure

- Power transformers → vital parts of every power network
- Winding failures → the predominant causes of failures
- High percentage of failures → insulation problems (about 41%)



Transformer failure for transformer at substation [1]



Failure statistics of power transformer component based failures [2]

[1] Hussain, Md Rashid, Shady S. Refaat, and Haitham Abu-Rub. "Overview and partial discharge analysis of power transformers: A literature review." *IEEE Access* 9 (2021).

[2] Murugan, Raji, and Raju Ramasamy. "Failure analysis of power transformer for effective maintenance planning in electric utilities." *Engineering Failure Analysis* 55 (2015).

Power Transformers Failure

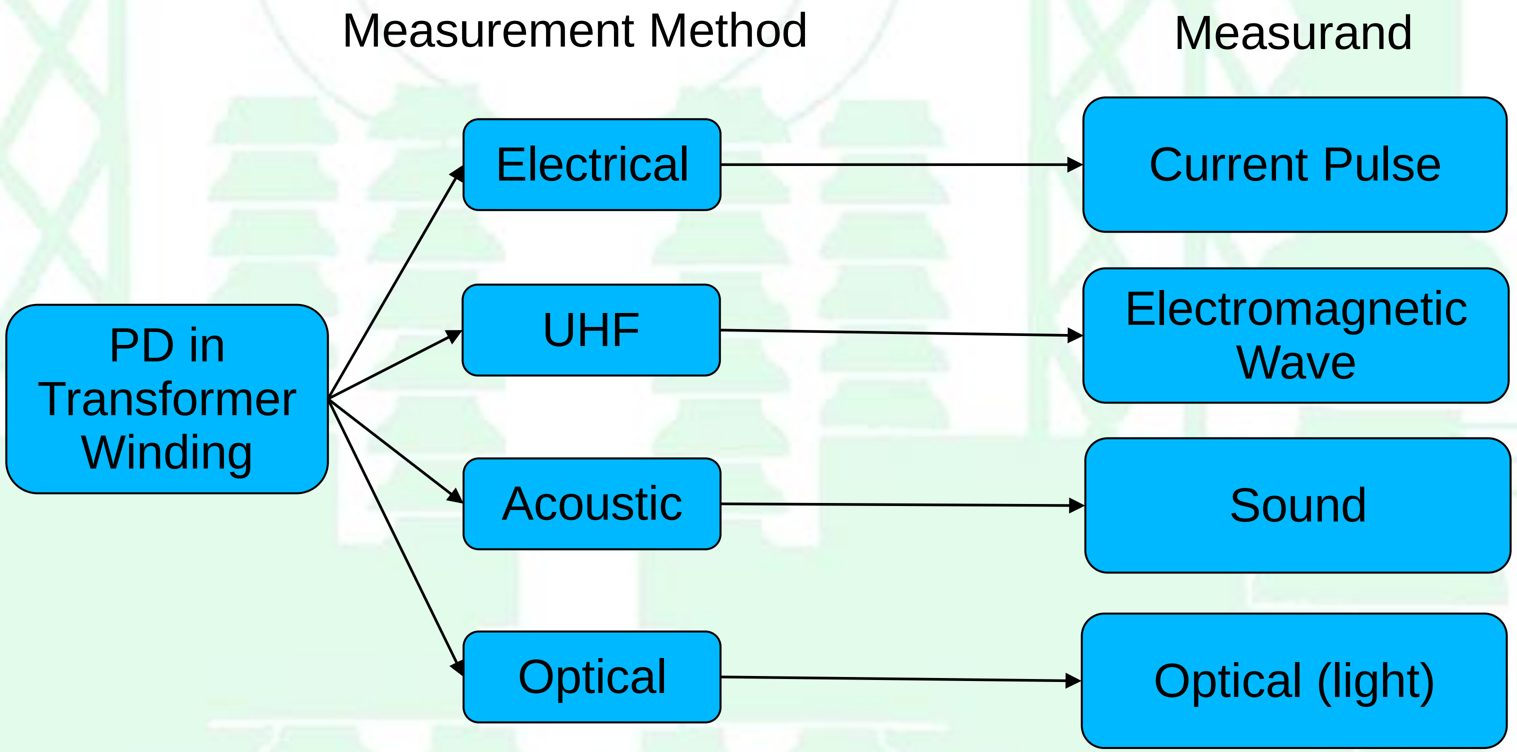
- Many winding problems can be detected by Partial Discharge (PD) monitoring and localization
- PD has several impacts in transformers
  - Accelerated degradation of insulation materials
  - Overheating due to high energy PD weakening the whole system
  - Reduced life expectancy of the transformer
  - Worst case scenario: unexpected breakdown

Bad insulation paper due to partial discharge

https://www.apolloenergyanalytics.com/transformer-failure-partial-discharge/

PD Localization in a Power Transformer Winding

https://www.electricalclassroom.com/types-of-transformer-windings/



PD Localization in a Power Transformer Winding

https://www.electricalclassroom.com/types-of-transformer-windings/

Measurement Method

PD in Transformer Winding

Electrical

UHF

Acoustic

Optical

Less prone to the medium noises

Easier to implement

More robust to attenuation

PD Localization Using Electrical Methods Procedures

Objective:
Simulate PD in a transformer winding and localize its source using a learning-based algorithm

PD Simulation

1. *Modelling of the Winding*
2. *Modelling of a PD Pulse*
3. *Injecting the PD Pulse to the Winding*
4. *Capturing the Output Signal at the Winding Terminal*

PD Localization

1. PD Data Acquisition
2. Feature Extraction for PD Signal
3. PD Data Analysis
4. PD Localization

PD Localization Using Electrical Methods

Modelling of the Winding

Modelling of a PD Pulse

Injecting the PD Pulse to the Winding

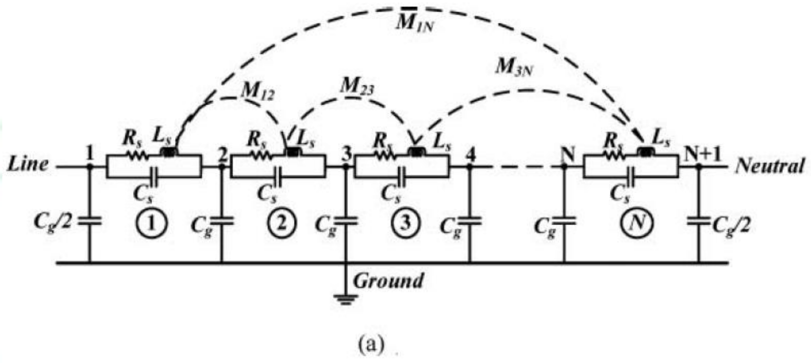
Capturing the Output Signal at the Winding Terminal

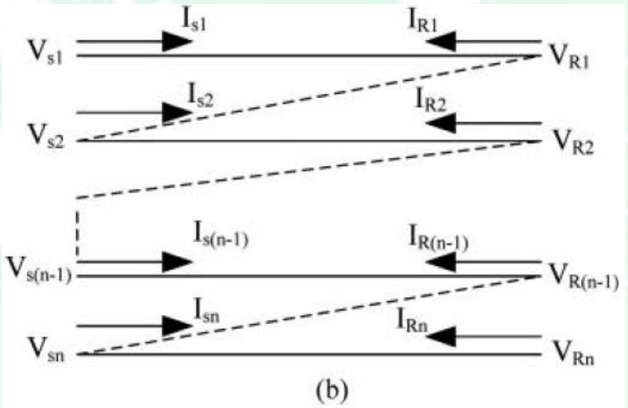
Valid for the low frequency range

- Capacitive network model
- Detailed ladder network model

Valid for the high frequency range

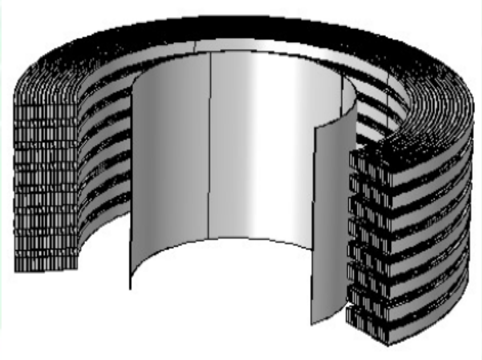
- Multi-conductor transmission line (MTL)
- Axial multi-conductor transmission line (AMTL)





(a) Detailed ladder network model and (b) multi-conductor transmission line model (MTL) [1]

3D split view of the winding



[1] M. Mondal and G. B. Kumbhar, “Partial discharge localization in a power transformer: methods, trends, and future research,” *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, vol. 34, no. 5. Taylor and Francis Ltd., Sep. 03, 2017.

PD Localization Using Electrical Methods

Modelling of the Winding

Modelling of a PD Pulse

Injecting the PD Pulse to the Winding

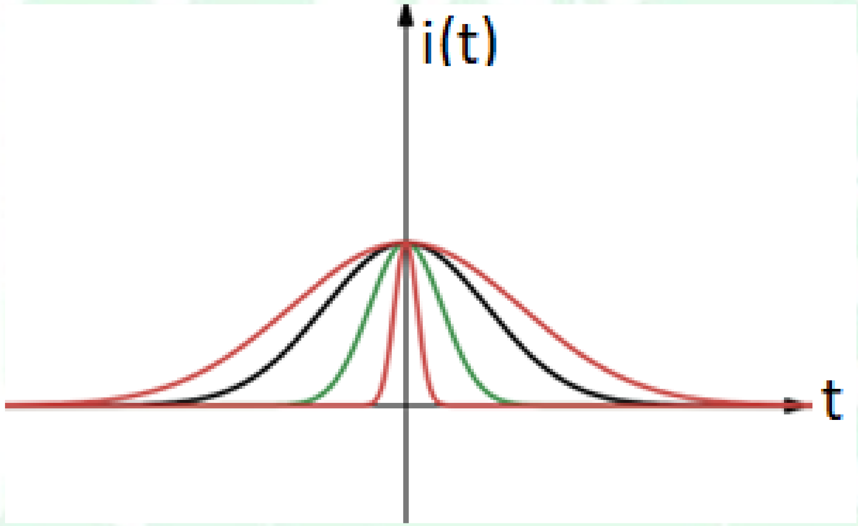
Capturing the Output Signal at the Winding Terminal

Single Gaussian pulse with a specified peak and rise time

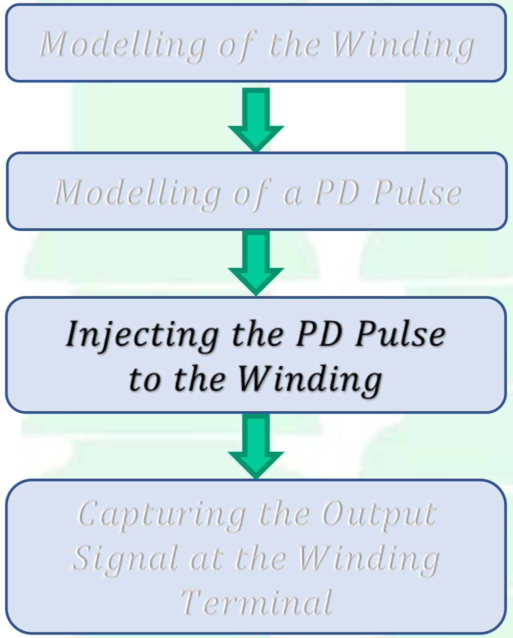
$$\boldsymbol{i(t) = Ae^{-(t-t_0)^2/\alpha^2}}$$

where
A → Amplitute
$\boldsymbol{t_0}$ → Time delay
α → Standard deviation

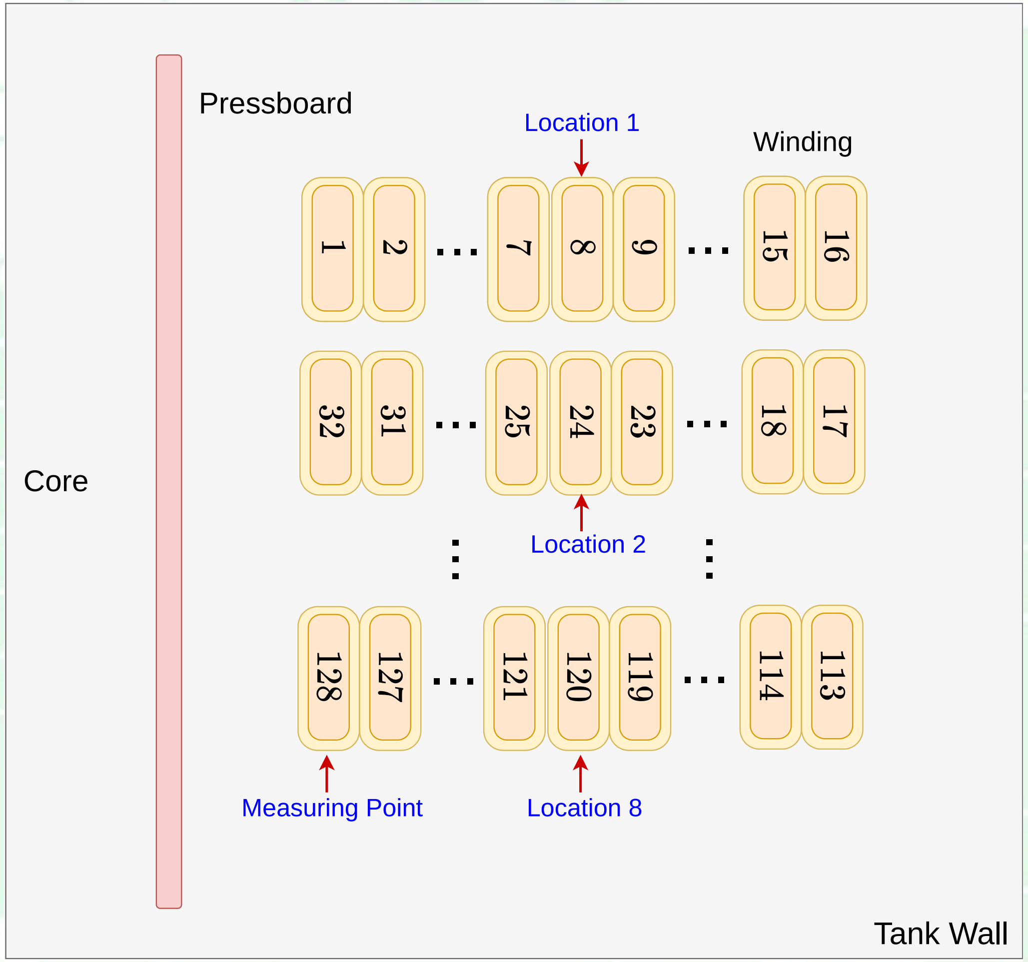
1ns < α < 20ns



PD Localization Using Electrical Methods

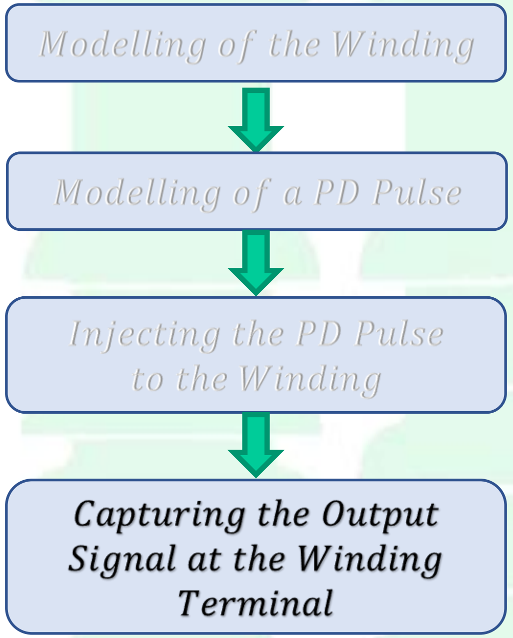


- An eight-disk winding with 128 turns
- The middle turn of each section → the injection locations
- The end of the last turn → grounded via a small resistance
- The other end → open circuit

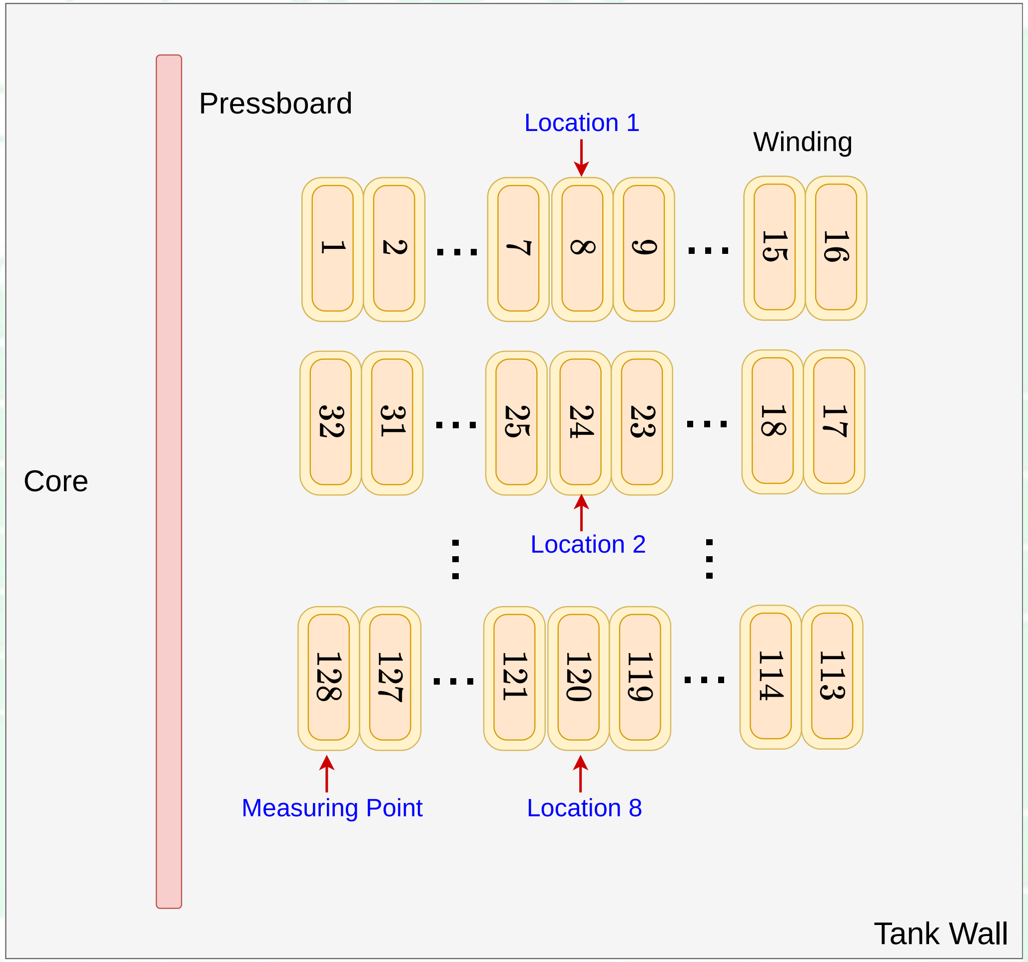


Schematic of 2D cross section of the winding

PD Localization Using Electrical Methods



- Injecting the Gaussian pulse to the locations → different rise times
- The current of the last turn → stored in each time iteration as an output



Schematic of 2D cross section of the winding

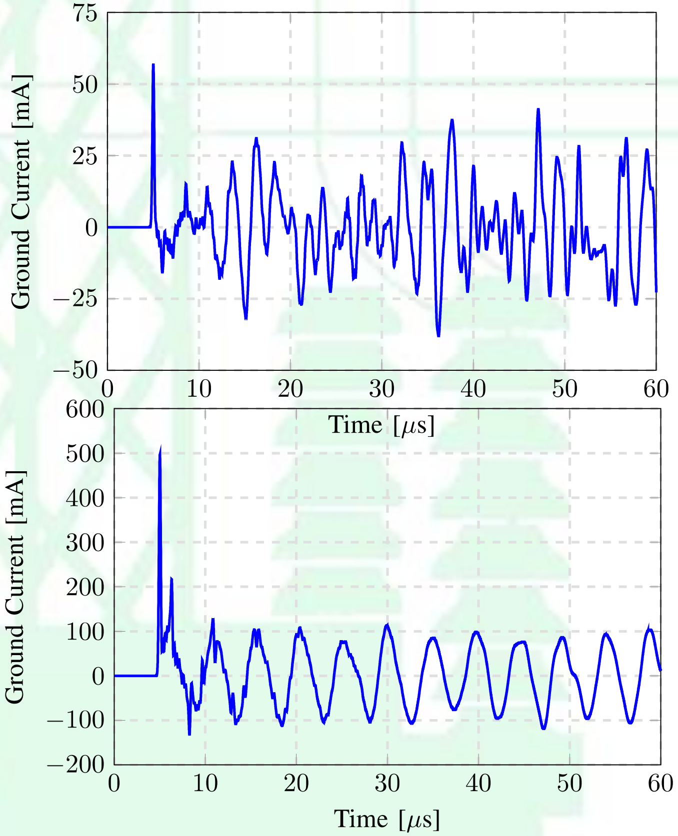
PD Localization Using Electrical Methods

Modelling of the Winding

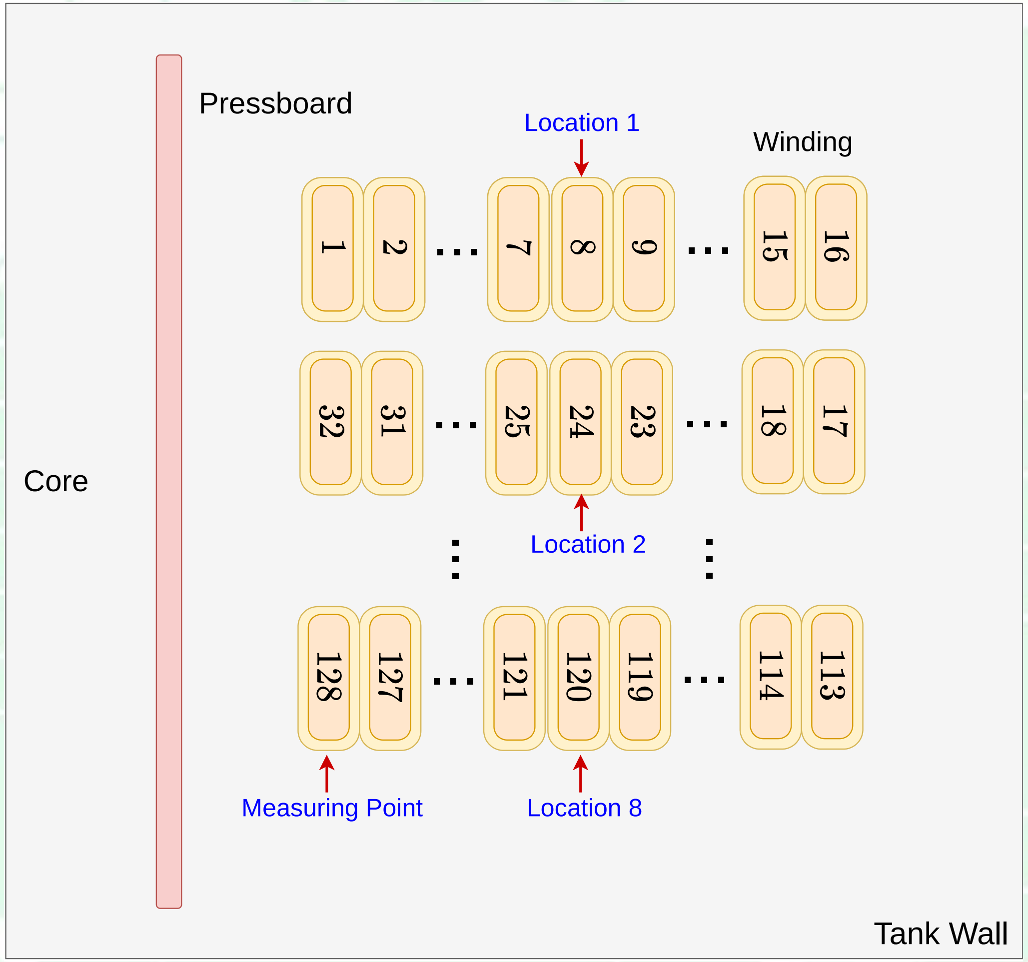
Modelling of a PD Pulse

Injecting the PD Pulse to the Winding

Capturing the Output Signal at the Winding Terminal



Two sample current waveforms recorded at the ground terminal when the PD pulse is injected at location 1 and location 8



Schematic of 2D cross section of the winding

PD Localization Using Electrical Methods

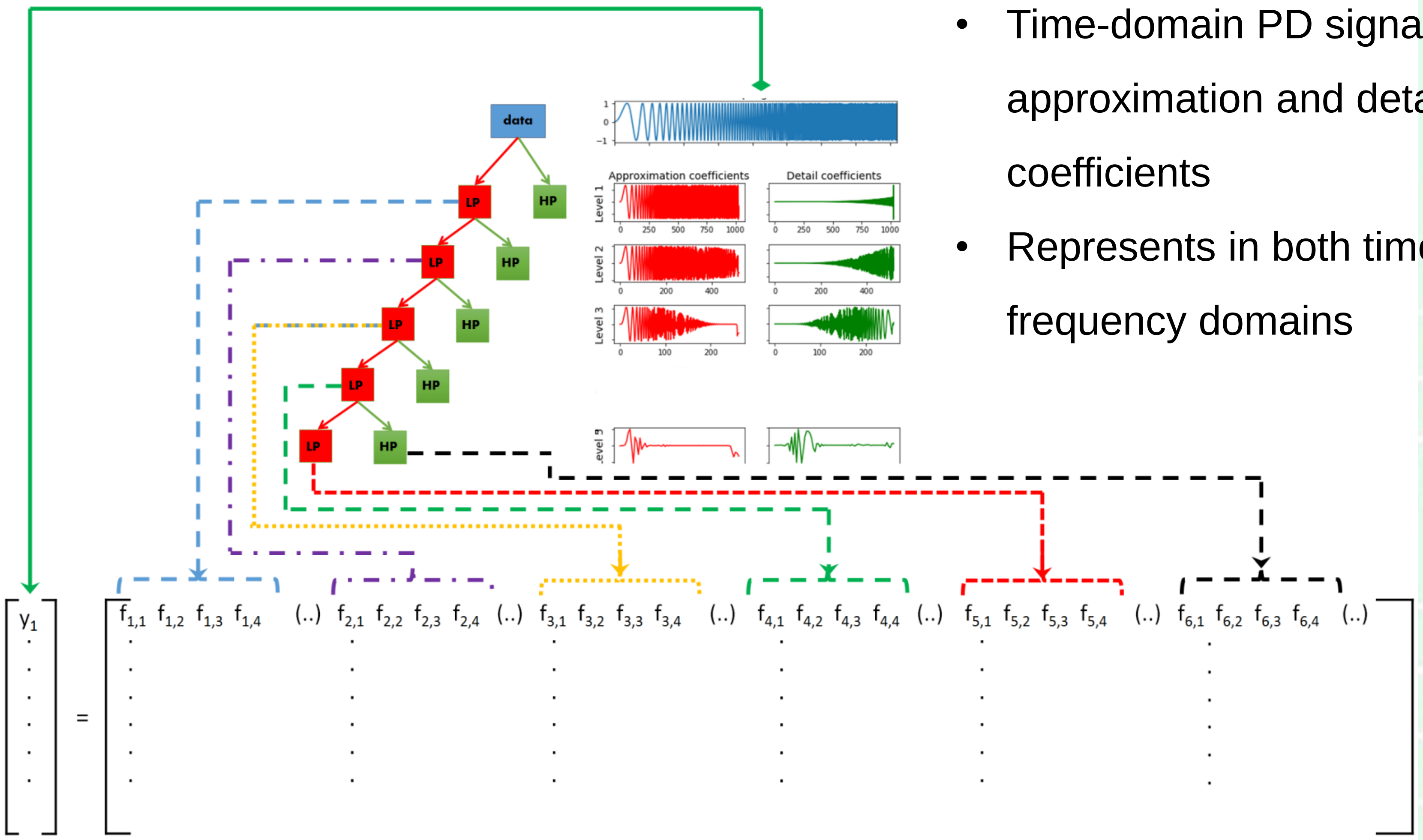
PD Data Acquisition and Preprocessing

Feature Extraction for PD Signal

PD Data Analysis

PD Localization

Wavelet Transform



- Time-domain PD signals → approximation and detail coefficients
- Represents in both time and frequency domains

https://ataspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning/

PD Localization Using Electrical Methods

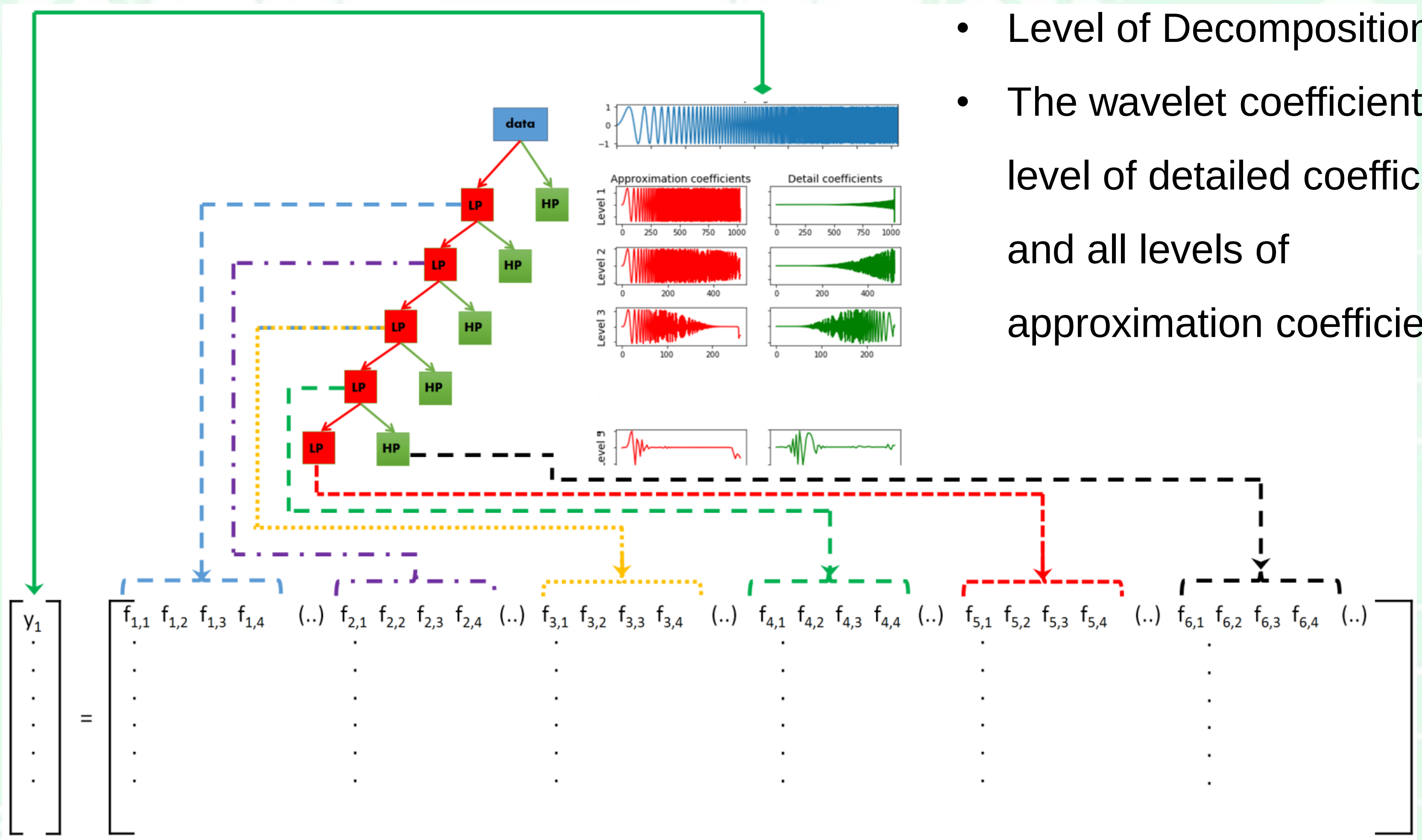
PD Data Acquisition and Preprocessing

Feature Extraction for PD Signal

PD Data Analysis

PD Localization

Wavelet Transform



- Level of Decomposition = 9
- The wavelet coefficients: Last level of detailed coefficients and all levels of approximation coefficients

https://ataspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning/

PD Localization Using Electrical Methods

PD Data Acquisition and Preprocessing

↓

Feature Extraction for PD Signal

↓

PD Data Analysis

↓

PD Localization

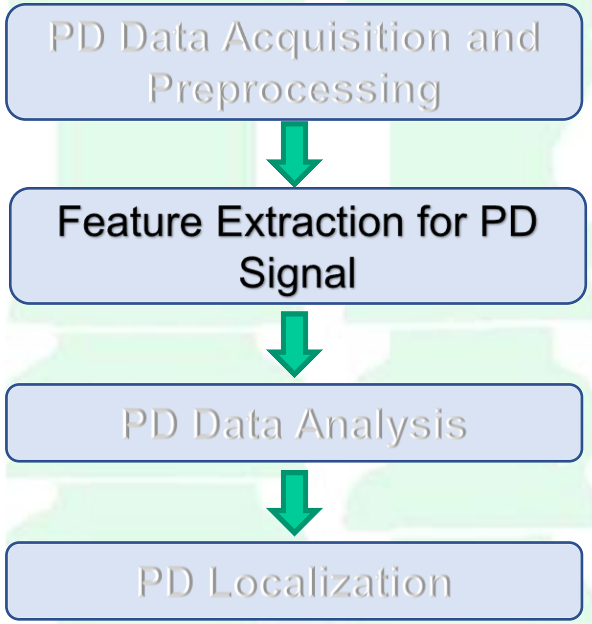
Hand-crafted Feature Extraction

- manually based on user experience and expertise
- Time-consuming
- No guarantee that the best features are selected

Automated Feature Extraction

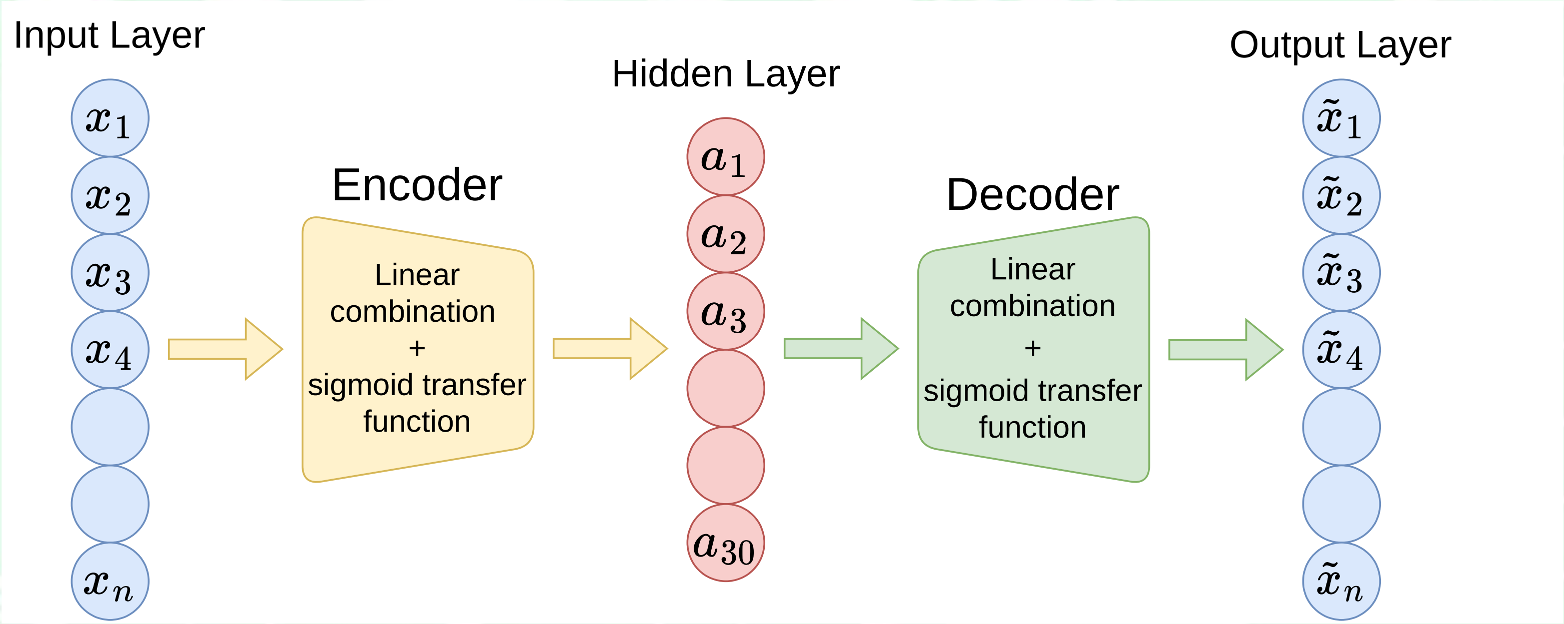
- Developing a learning model to learn the features
- Sparse Autoencoder (SAE) in this work

PD Localization Using Electrical Methods



Feature Extraction using SAE

- The number of input and output nodes are equal, and the model should learn to optimize their similarity
- The number of nodes in the hidden layer represents the number of features



Sparse Autoencoder architecture with *n* input/output nodes and 30 hidden nodes.

PD Localization Using Electrical Methods

PD Data Acquisition and Preprocessing

Feature Extraction for PD Signal

PD Data Analysis

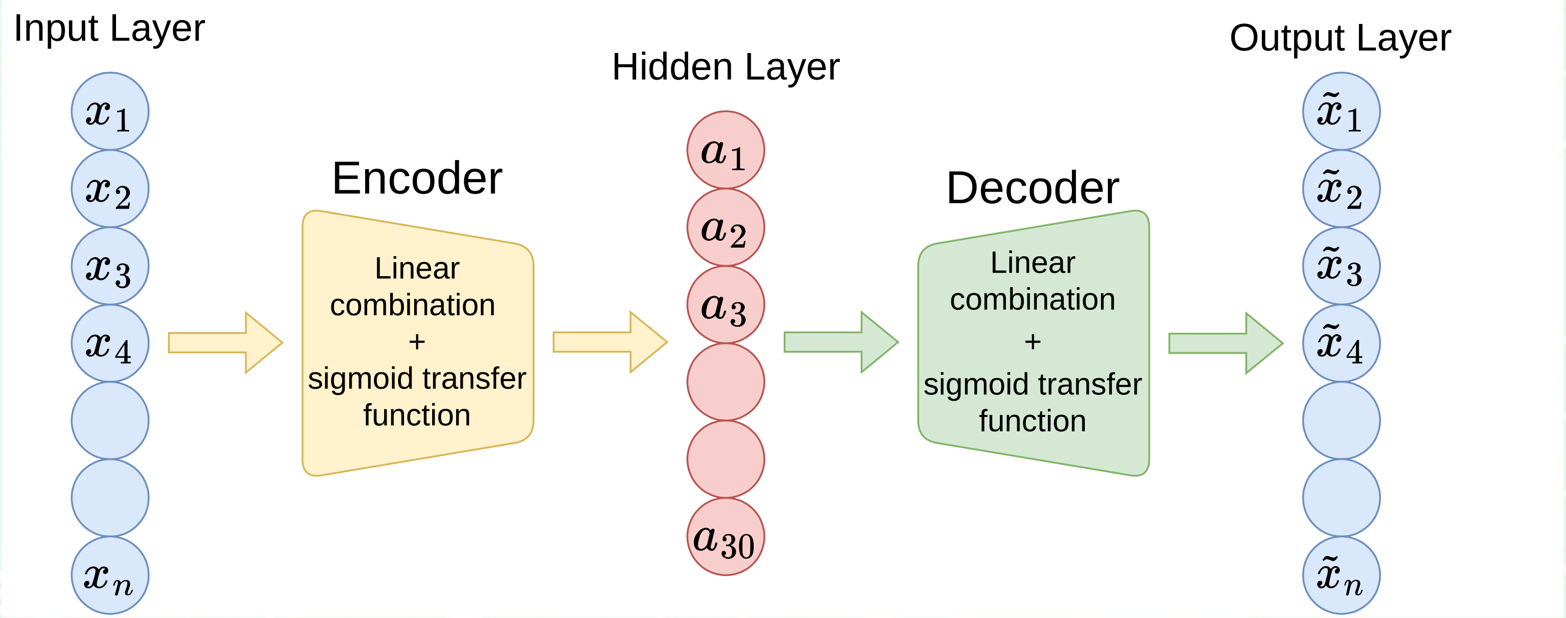
PD Localization

$$
\begin{aligned}
a_1 &= f(W_{11}^{(1)} x_1 + W_{12}^{(1)} x_2 + W_{13}^{(1)} x_3 + \cdots) \\
a_2 &= f(W_{21}^{(1)} x_1 + W_{22}^{(1)} x_2 + W_{23}^{(1)} x_3 + \cdots) \\
a_3 &= f(W_{31}^{(1)} x_1 + W_{32}^{(1)} x_2 + W_{33}^{(1)} x_3 + \cdots)
\end{aligned}
$$

$$f(z) = \frac{1}{1 + \exp(-z)}$$

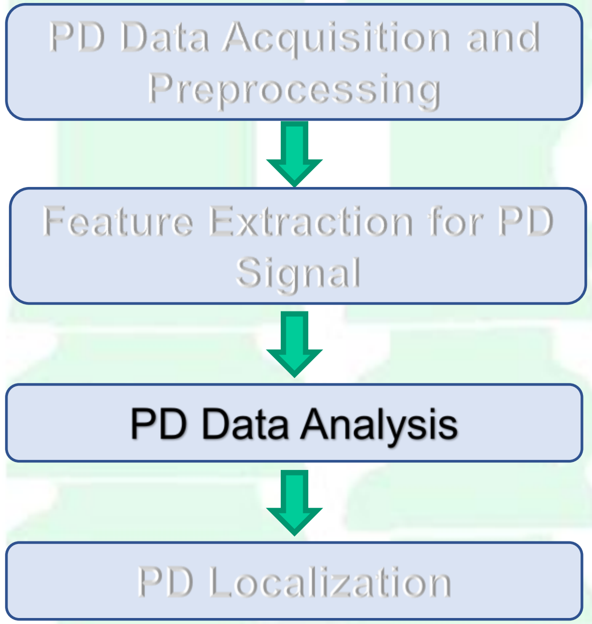
$$\tilde{x}_1 = f(W_{11}^{(2)} a_1 + W_{12}^{(2)} a_2 + W_{13}^{(2)} a_3 + \cdots)$$

$$\tilde{x} \approx x$$



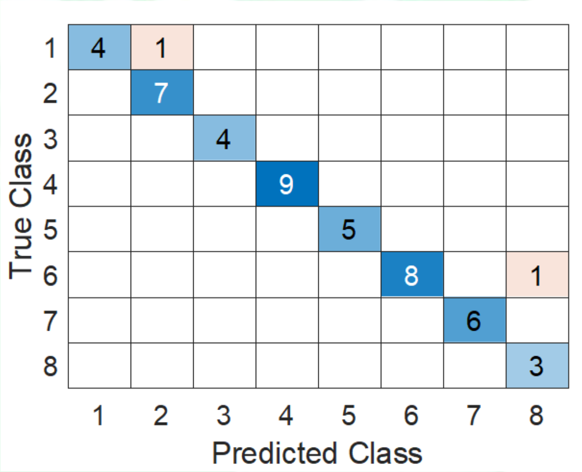
Sparse Autoencoder architecture with *n* input/output nodes and 30 hidden nodes.

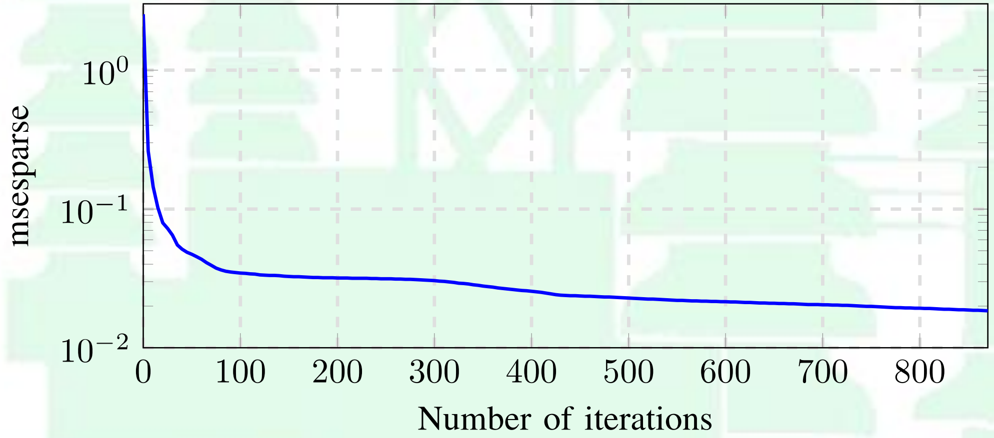
PD Localization Using Electrical Methods



Logistic regression, one-vs-all classifier

- 70% of the data → training set and 30% → test set
- Regularization = 0.001
- Training Accuracy: 99.1% Test Accuracy: 97.9%





a) The confusion matrix for the proposed model test set, b) mean squared error with L2 and sparsity regularizers (msesparse) vs the number of iterations.

PD Localization Using Electrical Methods

PD Data Acquisition and Preprocessing

↓

Feature Extraction for PD Signal

↓

PD Data Analysis

↓

PD Localization

- Repeating the algorithm using a different number of hyperparameters
- No improvement in test accuracy in any of these modified models

Table 1: The accuracy of the classifier with different value of hyperparameters.

	Training Accuracy (%)	Test Accuracy (%)
The proposed model	99.1	97.9
Level of decomposition = 8	99.1	93.8
Regularization = 0	100	93.8
Regularization = 0.0001	95.5	93.8
Number of Features = 29	97.3	91.7
Number of Features = 31	98.2	93.8

PD Localization Using Electrical Methods

PD Data Acquisition and Preprocessing

⇓

Feature Extraction for PD Signal

⇓

PD Data Analysis

⇓

PD Localization

- Extracting the features manually and using as input features
- Significant reduction in the performance in all the cases
- Worse performances un the other possible combinations of features in terms of test accuracy.

Table 2: The accuracy of the classifier with different value of hyperparameters with hand-crafted feature extraction.

	Training Accuracy (%)	Test Accuracy (%)
Statistical features only	86.8	78.3
Energy and statistical features	88.7	73.9
Entropy and statistical features	97.2	69.6
All features together	97.2	71.7

Conclusion

- A transmission-line-based model was used to simulate an eight-disk transformer winding.
- The SAE model was employed with different hidden nodes to determine the optimal number of features.
- A logistic regression, one-vs-all classifier, was then employed to localize PD in the winding.
- The classification results showed an accuracy of 99.1% for the training set and 97.9% for the test set.
- The presented method was repeated using different hyperparameters, but no improvement was seen.
- The classification was conducted without the SAE using features recommended in previous literature.
- The comparison showed a significant reduction in the performance of the classifier, which indicates the performance enhancement of the automated feature extraction over handcrafted feature extraction.

Financial support from University of Manitoba and NSERC is acknowledged.