



GPT-like Attention
Mechanisms for Power
Transformer Condition
Monitoring and Prognostics

Mentor: Yufei Tang, Ph.D.

Raul Mendy, Aira Torres,
Joseph Accurso

Research Objectives

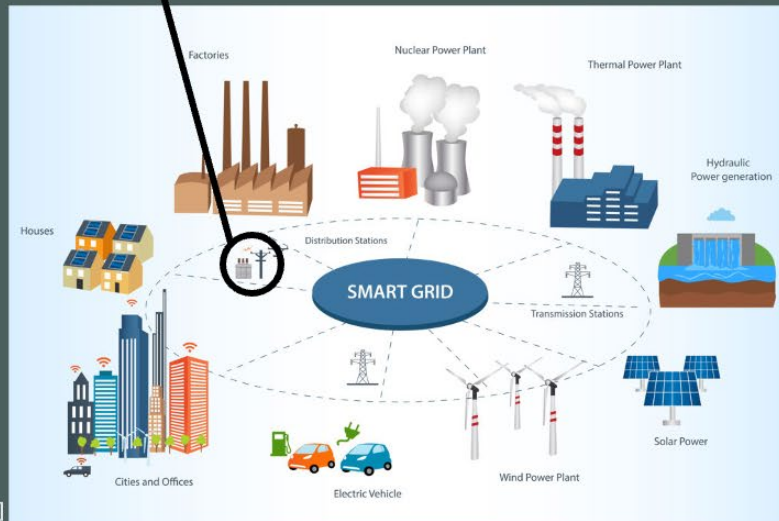


Power Failure Indication

Reactive Maintenance

Scheduled Maintenance

On-Site System Diagnostics



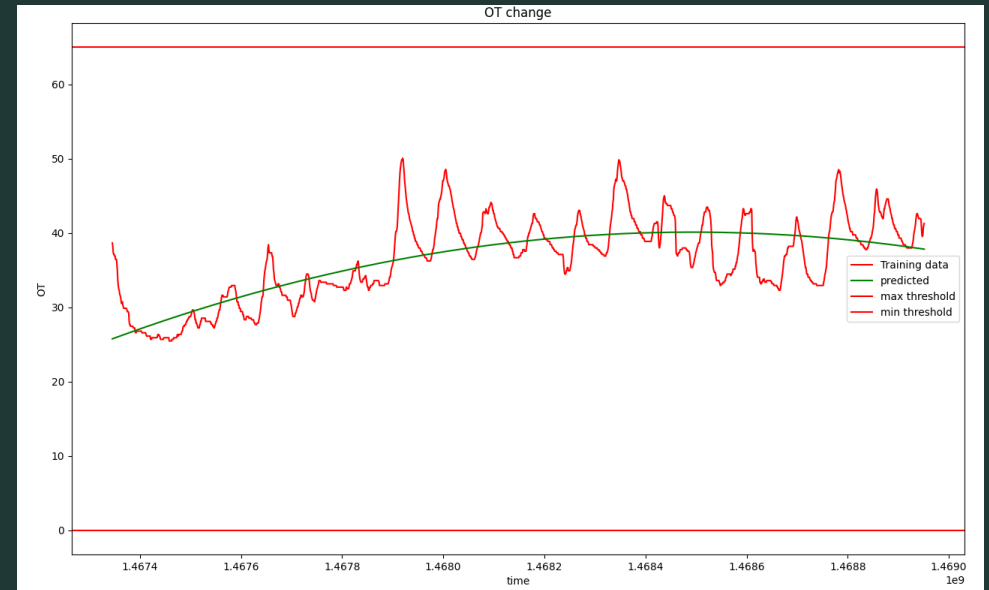
REMAINING USEFUL LIFE MODEL

Remaining Useful Life

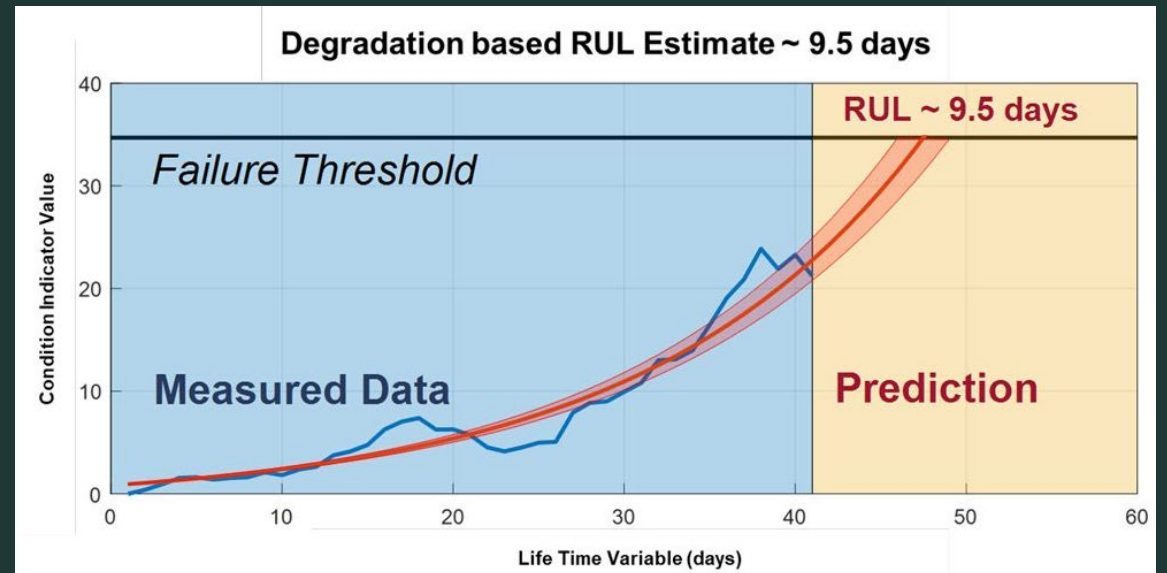
- ETTm1 & ETTm2
- 2 power transformers
- Change in oil temp for 2 years
- Degradation-based RUL model
- Min and max threshold values
- Estimates # of days until maintenance is required

Machine Learning Model:

- Polynomial Regression



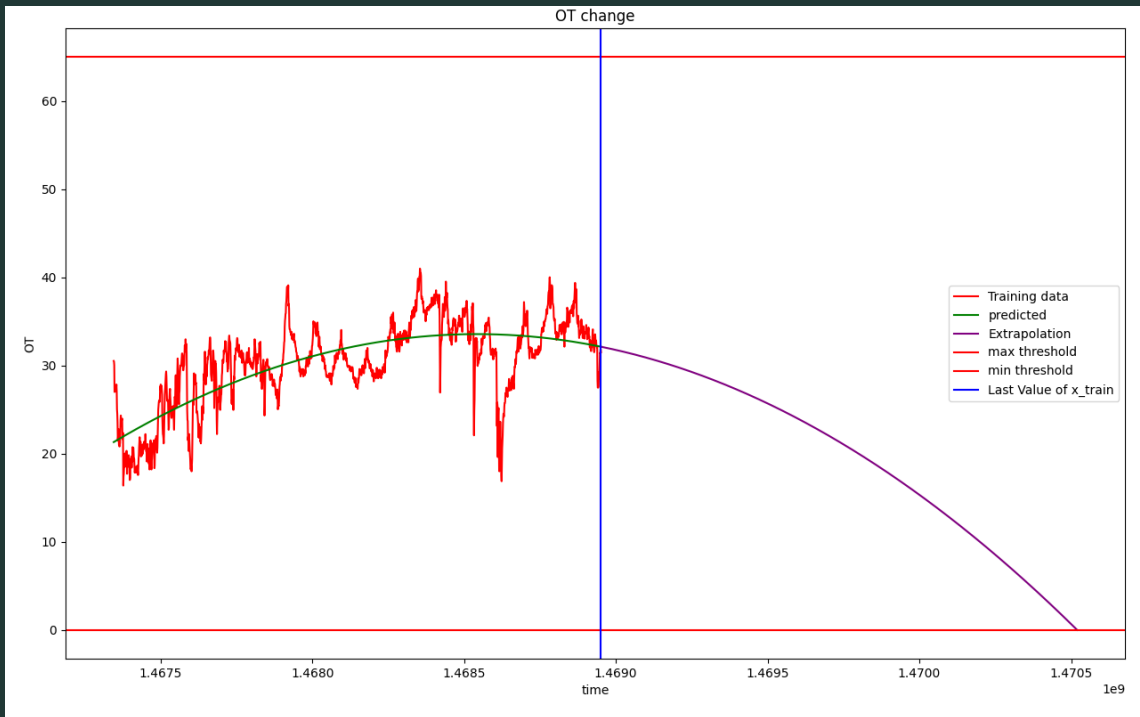
ETTm2:MONTH 1



Remaining Useful Life Results

Case 1: Decreasing prediction line

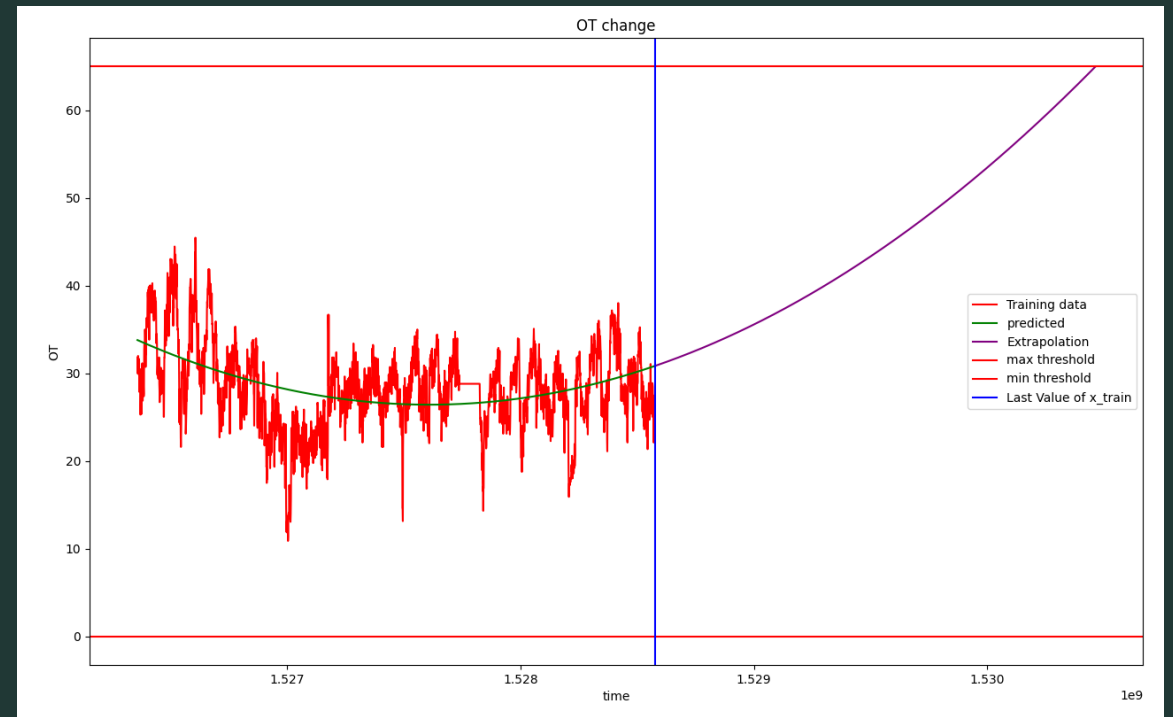
Month 1: M1 dataset



RUL = 18 DAYS

Case 2: Increasing prediction line

Month 24: M2 dataset

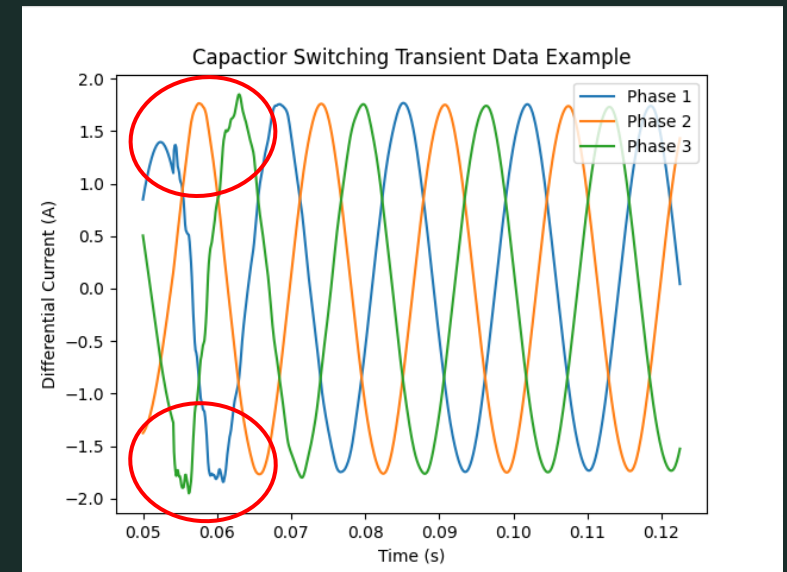
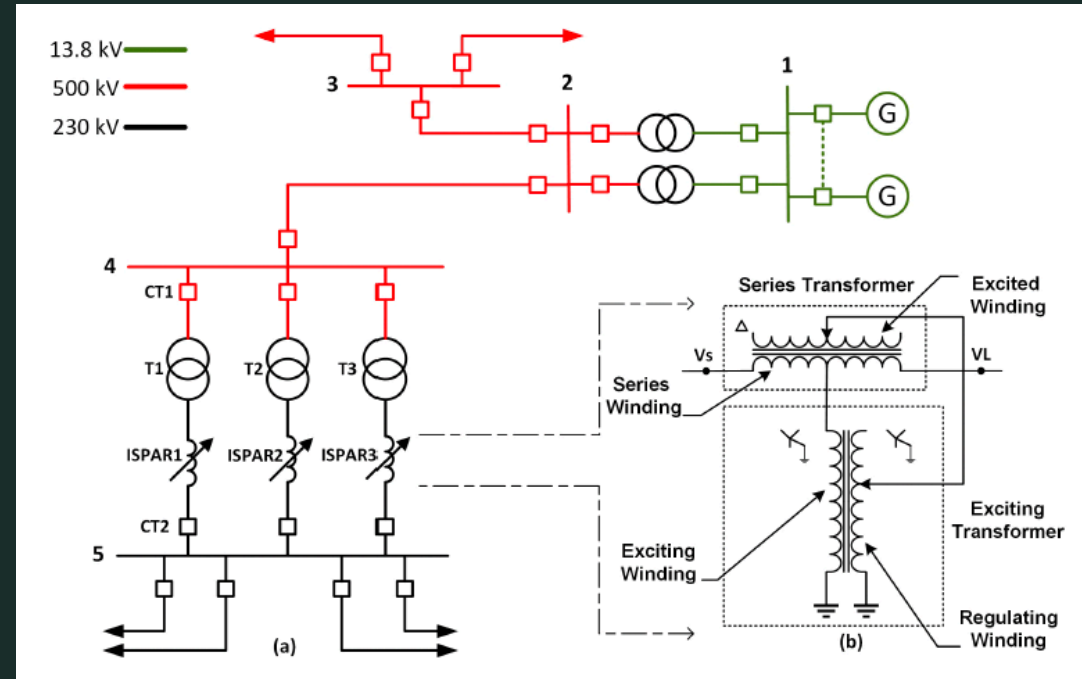


RUL = 22 DAYS

MULTI -CLASS CLASSIFICATION MODEL

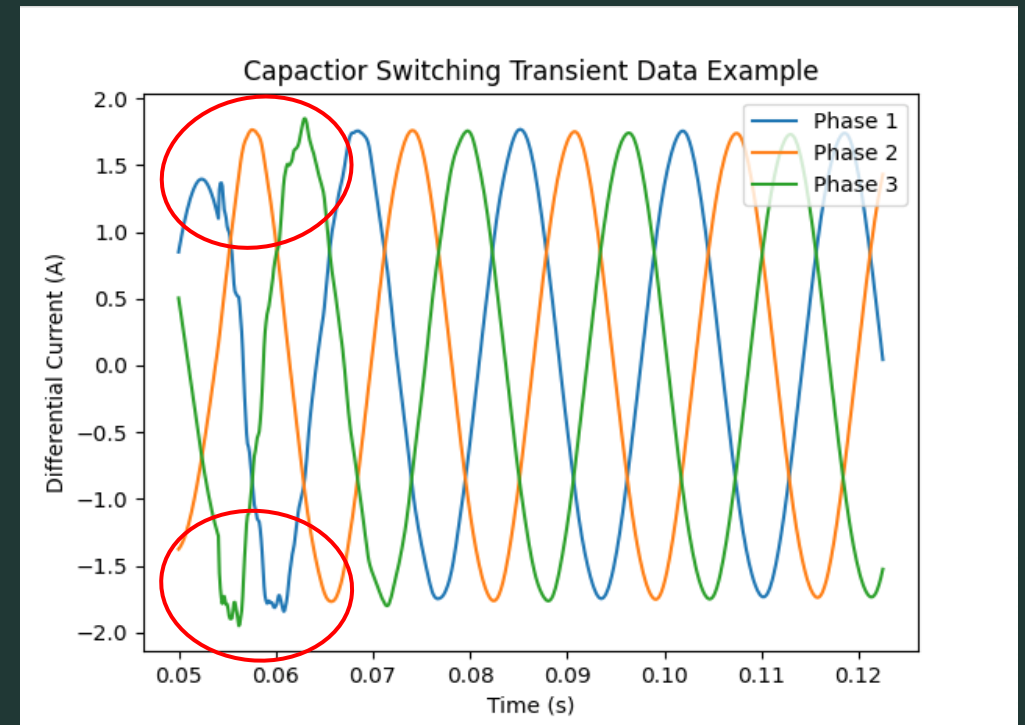
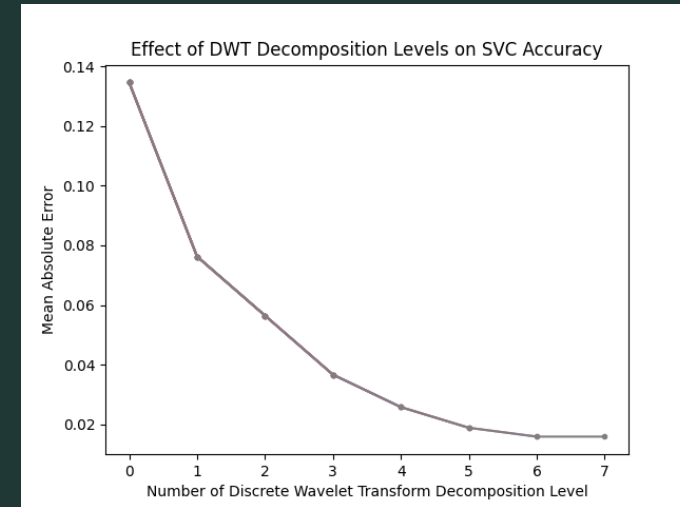
Fault Classification Dataset

- Text File Format
- ~110,000 Simulated Examples
- 45 Fault Classes
- 3 Phases
- 726 Measurements per Phase
- Normalized to ~390,000 Examples
- ~850,000,000 Total Measurements



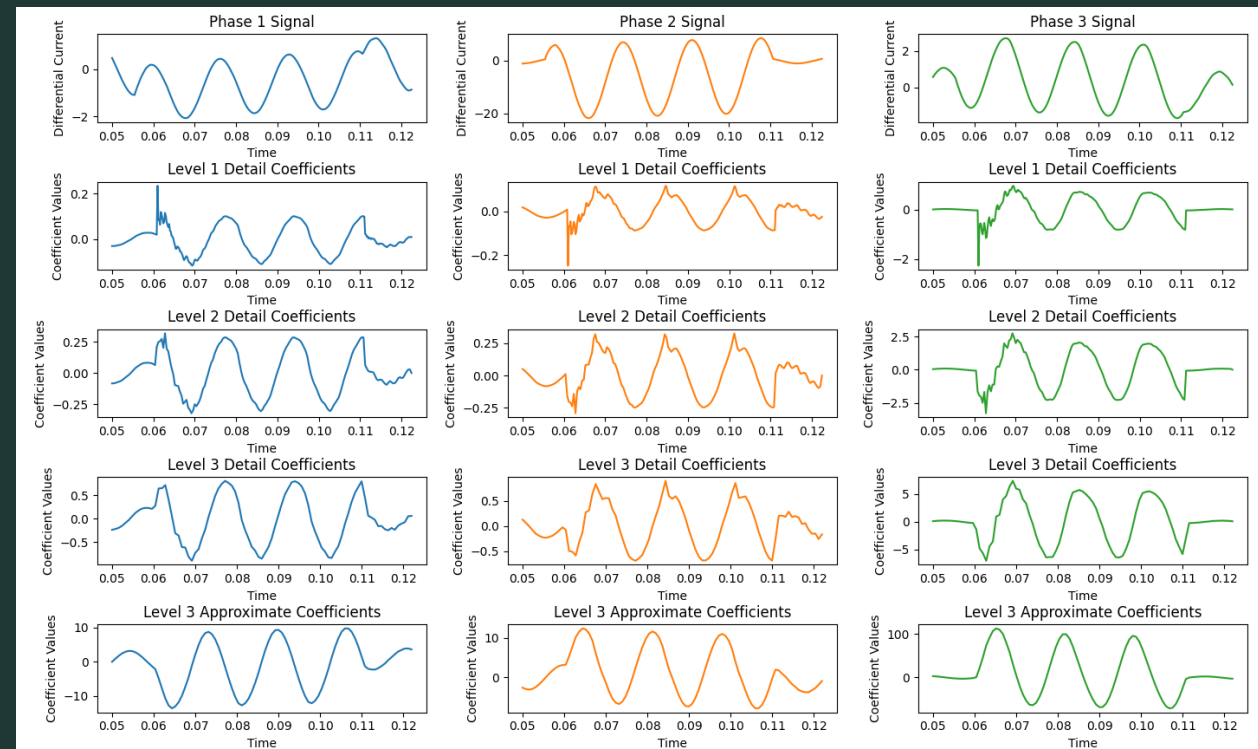
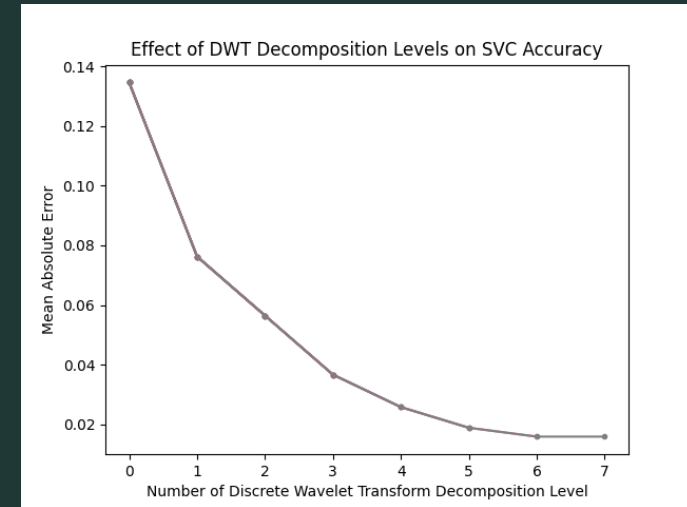
Fault Classification Using Machine Learning

- Discrete Wavelet Transform Decomposition
 - Daubechies 2 Wavelet
 - 3 - 5 Decomposition Levels
- Random Forest
- Gradient Boost
- Support Vector Machine



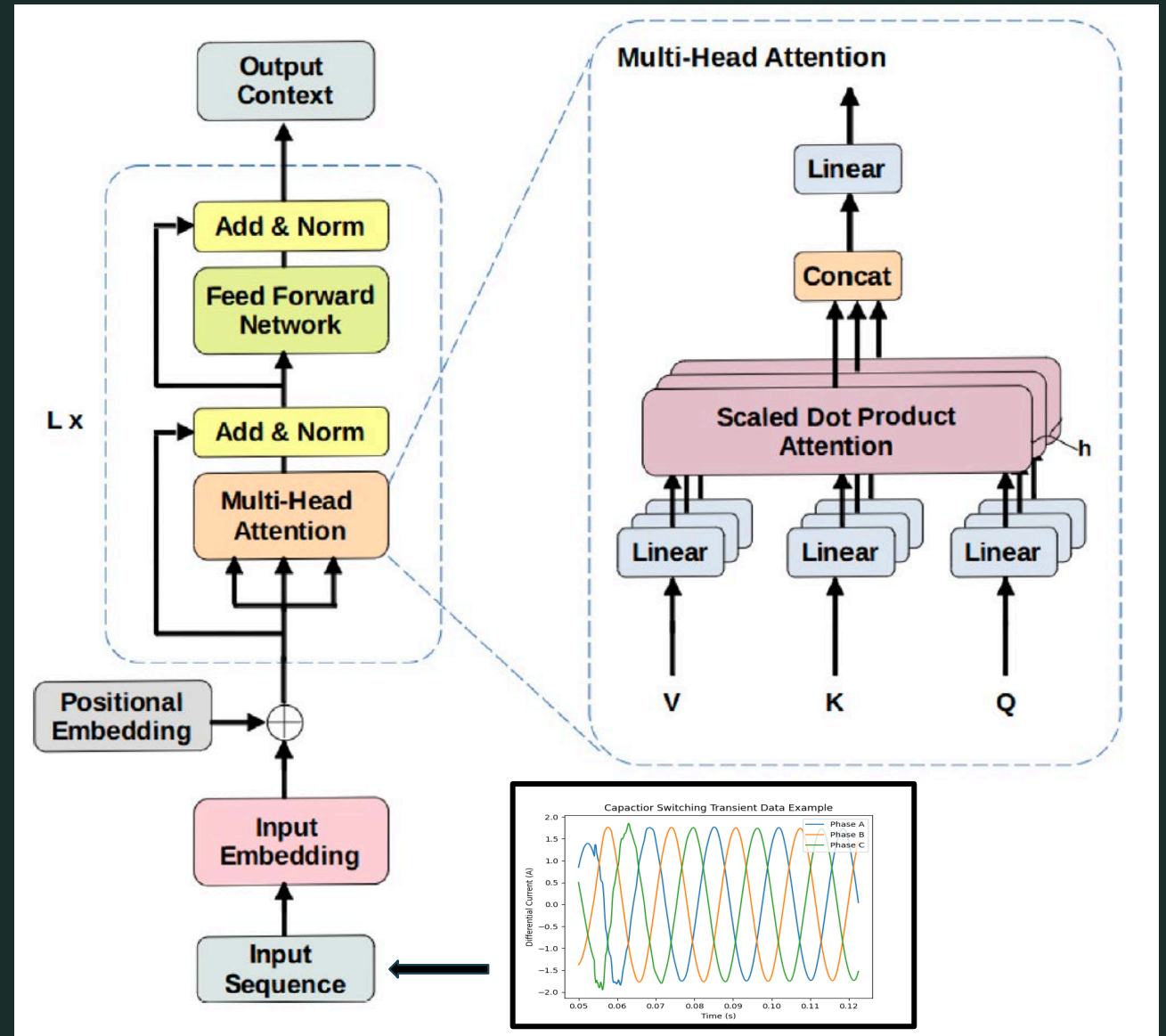
Fault Classification Using Machine Learning

- Discrete Wavelet Transform Decomposition
 - Daubechies 2 Wavelet
 - 3 - 5 Decomposition Levels
- Random Forest
- Gradient Boost
- Support Vector Machine



Fault Classification Using the Transformer Architecture

- Foundational architecture of Chat-GPT
- Signal analysis within positional context
- Multi-head attention mechanism
- No need for time-consuming DWT decomposition-based feature extraction



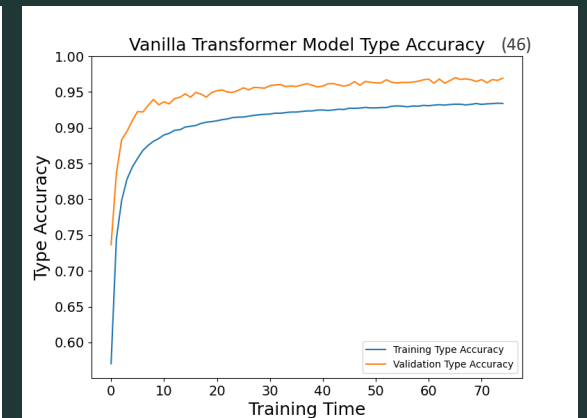
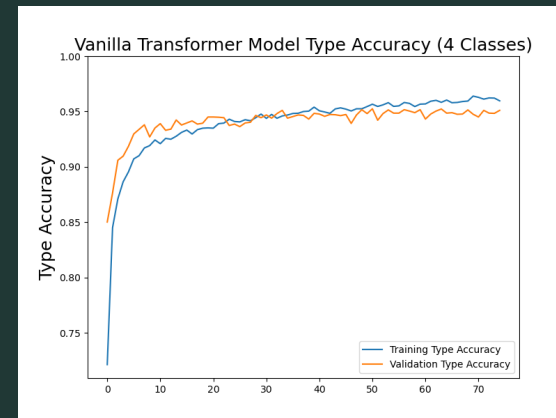
Fault Classification Results

Machine Learning Models

Model Type	# Classes	Accuracy	Training Time
Random Forest	2	92%	7 minutes
Gradient Boost	2	91%	20 minutes
SVC	2	98%	5 minutes
SVC	4	92%	2 hours

Transformer Architecture

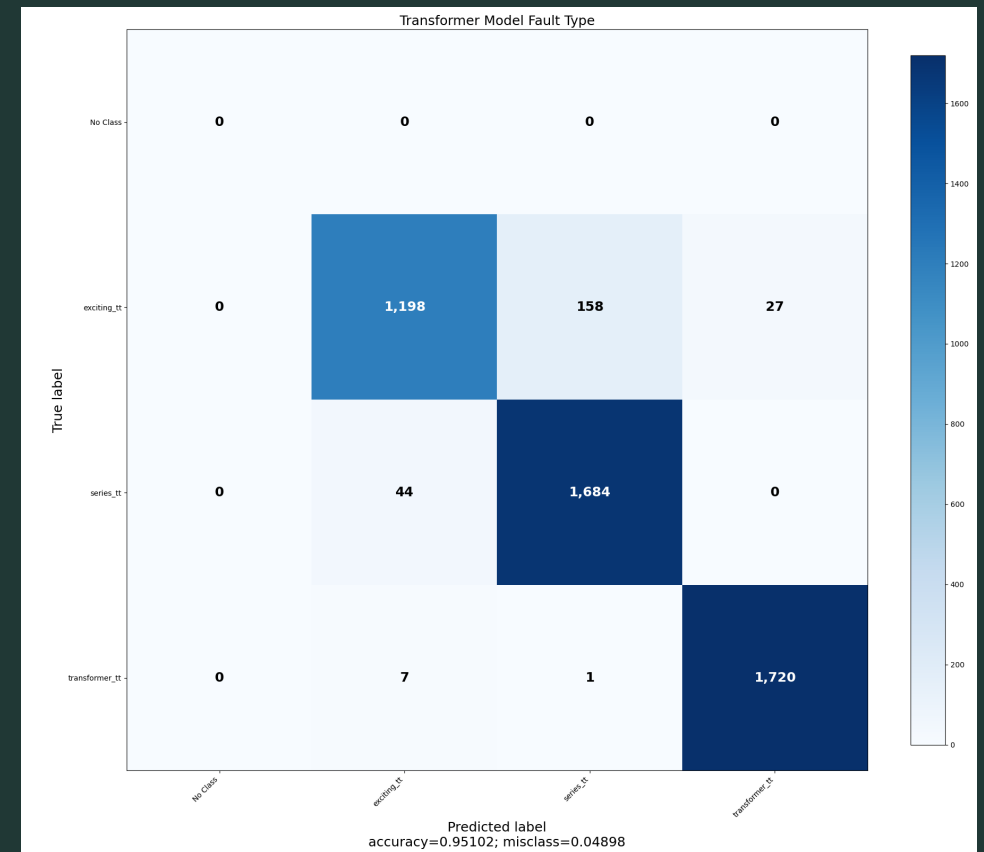
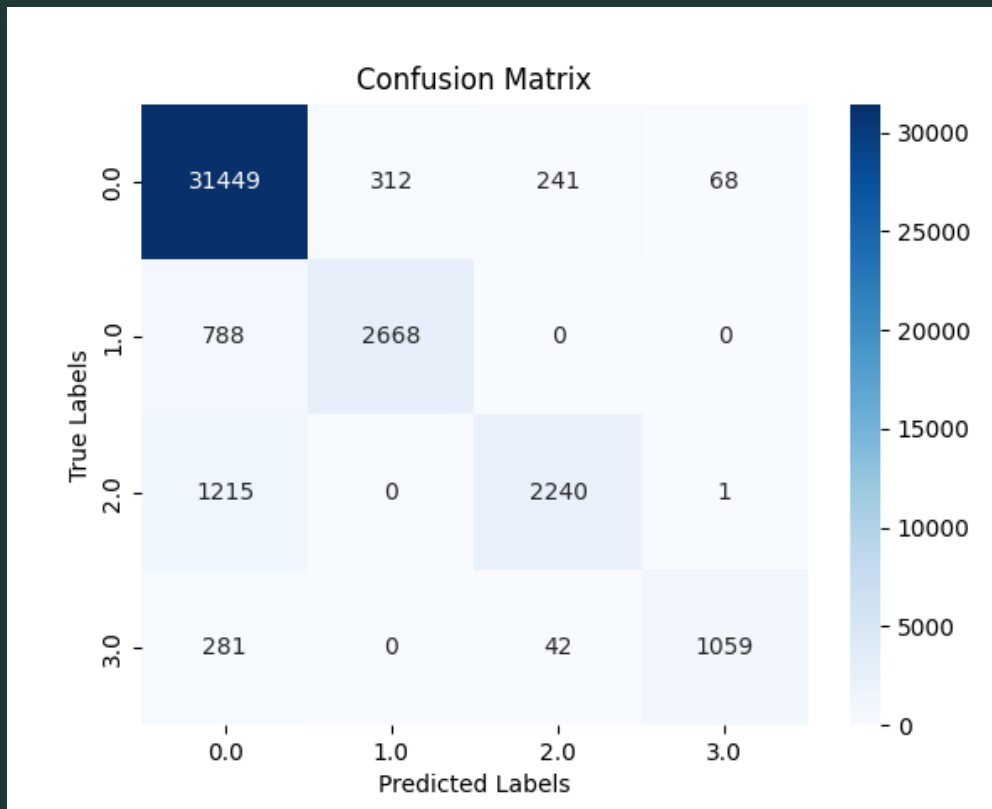
Model Type	# Classes	Accuracy	Training Time
Transformer Arch.	4	95%	30 minutes
Transformer Arch.	46	97%	5 hours



Confusion Matrices

Support Vector Classification

Transformer Architecture



FULL SYSTEM INTEGRATION



Oil Temperature

Anomalous Three-Phase Signal

Fault Classification Model

Remaining Useful Life (RUL) Estimation Model

Historical Database

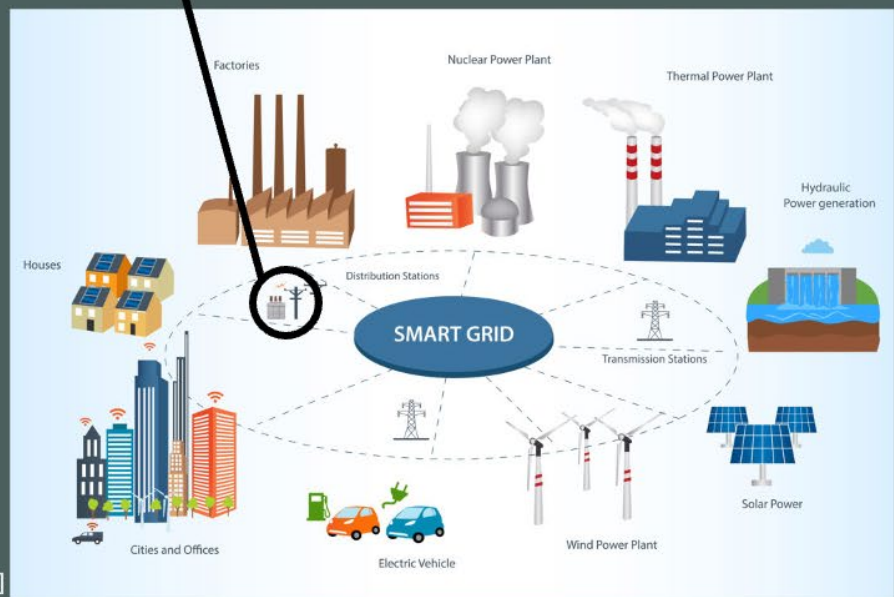
User Query

Natural Language Processor + Graphical User Interface

Natural Language Response

Diagnostic Information

Prognostic Estimation



DEMONSTRATION

FPL CHATBOT



[Large black area for chat messages]

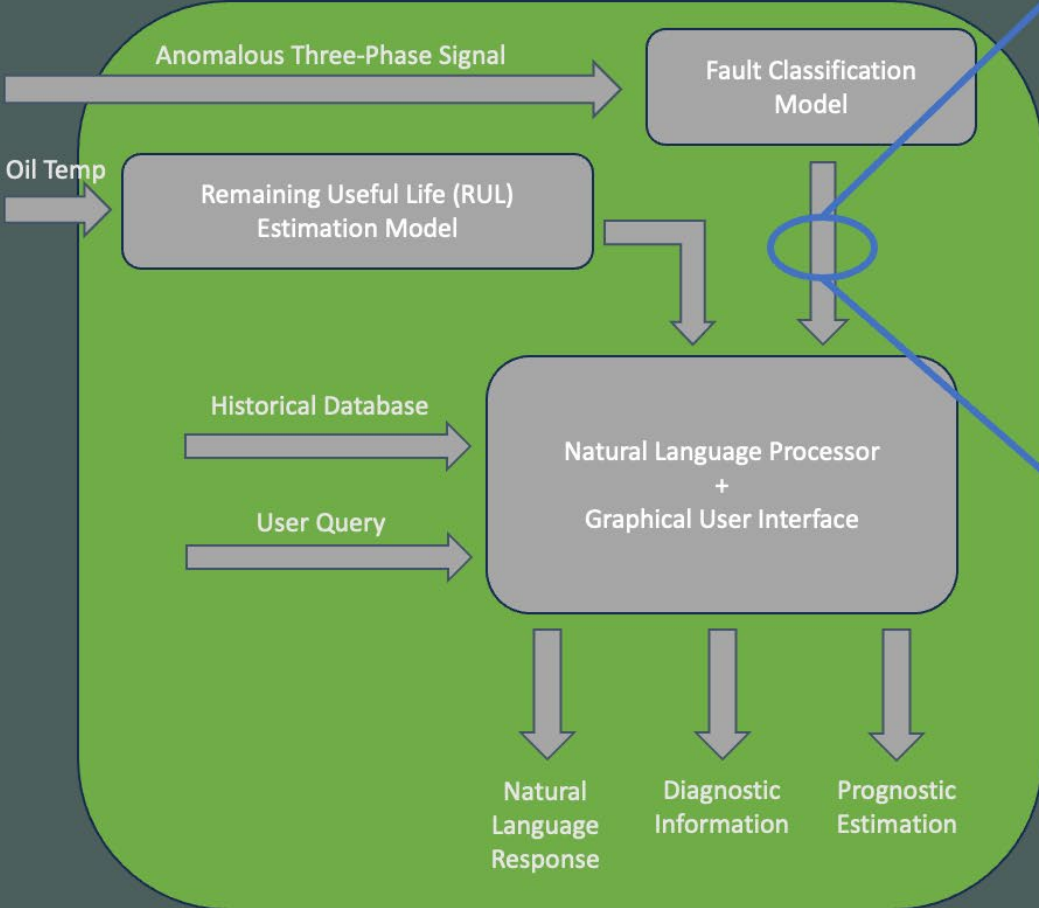
I

Send

Clear Screen

Next Steps

- FPL Data
- Duplicate results with real smart grid data
 - Use continuous data sources



Differential Architecture Search (DARTS) Transformer

Dynamically constructs the most optimal transformer architecture

The diagram shows four different transformer architectures, each with four layers labeled 0, 1, 2, and 3. The connections between layers are represented by colored lines (red, green, blue, yellow). The first architecture shows a sparse set of connections with question marks, representing the search space. The subsequent three architectures show increasingly dense and complex connection patterns, representing the search for an optimal architecture.

REFERENCES

- [1] Innovation at Work. (n.d.). The Smart Grid Could Hold the Keys to Electric Vehicles. IEEE. Retrieved from <https://innovationatwork.ieee.org/the-smart-grid-could-hold-the-keys-to-electric-vehicles/>
- [2] <https://tenor.com/view/transformer-gif-24446970>
- [3] Li, Y., Wang, H., & Sun, M. (2023). ChatGPT-Like Large-Scale Foundation Models for Prognostics and Health Management: A Survey and Roadmaps. *ArXiv*. /abs/2305.06472
- [4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *ArXiv*. /abs/1706.03762
- [5] Thomas, J. B., & K.V., S. (2023). Neural architecture search algorithm to optimize deep Transformer model for fault detection in electrical power distribution systems. *Engineering Applications of Artificial Intelligence*, 120, 105890. <https://doi.org/10.1016/j.engappai.2023.105890>
- [6] H. Ma, T. K. Saha, C. Ekanayake and D. Martin, "Smart Transformer for Smart Grid—Intelligent Framework and Techniques for Power Transformer Asset Management," in *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 1026-1034, March 2015, doi: 10.1109/TSG.2014.2384501
- [7] Pallav K Bera, Can Isik, Vajendra Kumar. (2020). Transients and Faults in Power Transformers and Phase Angle Regulators (DATASET). IEEE Dataport. <https://dx.doi.org/10.21227/1d1w-q940>