A ChatGPT-like Solution for Power Transformer Condition Monitoring

Joseph Accurso, Raul Mendy, Aira Torres, and Yufei Tang

Department of Electric Engineering and Computer Science
FPL Center for Intelligent Energy Technologies
Florida Atlantic University
Boca Raton, FL
Email: accu3879@ravens.benedictine.edu, {rmendy2019, airatorres2020, tangy}@fau.edu

***Abstract*—This paper explores the application of large-scale foundation models (LSF-Models) for prognostic health management of smart grid power transformers. One of the foundational architectures of many LSF-Models, including ChatGPT, is called the Transformer architecture. It enables LSF-Models to perform complex classification predictions and to provide the user with an appropriate web-based response. The advanced natural language processing techniques of the Transformer architecture interpret the nuances of power transformer-related data and efficiently extract valuable insights for fault detection and diagnosis. The multiple self-attention mechanisms of the architecture provide an enhanced form of feature extraction which expedites the creation of comprehensive models for large multiclass datasets. This fault classification method performed with 97.2% accuracy according to the Matthews Correlation Coefficient when evaluating 46 classes of simulated power transformer internal faults and transient disturbances. To better represent the diverse multimodal data present in smart grid systems, this research explores a unified system that integrates multiple machine-learning models into one simple and easy-to-use interface. The backend of this interface again takes advantage of the Transformer architecture to perform conversation-based classification with high accuracy and to provide a response or prediction from the appropriate system-integrated model. The success of the Transformer architecture throughout diverse applications within the same system showcases its potential to analyze a wide range of data typically found throughout a robust smart grid system.**

***Index Terms*—Large-Scale Foundation Models, Prognostic Health Management, Power System Transformer, Transformer Architecture, Multiclass Classification, Fault Detection**

I. Introduction

In 2022, the United States utilized 4.05 trillion kilowatt-hours of electricity to support residential, commercial, and industrial sectors [1]. With new emerging technologies connecting to the grid every day, the variety of energy consumption has put a diverse strain on the electrical grid system. The transmission of appropriate electrical energy from current distribution systems is only possible due to properly maintained electrical power transformers. These devices step up or step down voltage according to particular physical properties that take advantage of electromagnetic induction. This process, however, dissipates energy to the environment in the form of heat, thereby exposing the power transformer to potential thermal malfunctions. The general wear and tear built up throughout the lifespan of a power transformer contributes to additional transient disturbances and mechanical failures. Not only does this compromise the mechanical design of the power transformer but it also affects the stability of its electrical output. Anomalous voltage drops, voltage rises, and phase shifts all impact the overall reliability of the electrical grid and can be potentially damaging to any connected devices. Therefore, a power transformer operating with untreated mechanical faults becomes a weak link that, if left untreated, ultimately disrupts the entire chain of power distribution to industrial, commercial, and residential sectors.

The rise in electrical power dependence requires an even greater improvement in reliability. Due to the fact that internal faults and transient disturbances also impact the electrical signal itself, modern sensors incorporated into smart grid systems can be used to collect signal data and alert maintenance crews to defective power transformers. Although contemporary systems have shown promising results in detecting faults [2], classifying those detections poses a greater computational challenge. The amount of data needed to construct a comprehensive fault classification model for an entire smart grid system is vast and requires machine-learning models with the capability of handling extensive datasets. This is where large-scale foundation models (LSF-Models) such as GPT-3.5 (ChatGPT) [3] and Segment Anything Model (SAM) [4] have seen the greatest success. Applying an LSF-Model for smart grid system condition monitoring is a comprehensive solution that provides corroborated information through multiple models and simulations.

The success of many LSF-Models is attributed to the innovative Transformer deep learning architecture [5]. The Transformer architecture has become a popular foundation for many natural language processing (NLP) and computer vision (CV) models in the few years since its infancy in 2017. Its recent success in the form of ChatGPT proves its capability in handling large multiclass classification problems. This research proposes to use the Transformer architecture to classify three-phase signals in smart grids and to provide a similar conversational prediction to that of ChatGPT. Instead of processing global web data, the model was trained on anomalous time-series signals to provide a fault classification

This work was supported in part by the US National Science Foundation (NSF) through Grants XXX and XXX.

prediction. This multiclass fault classification model has the potential to be integrated with a basic natural language processor which also uses a simplified structure of the Transformer architecture to constitute the backend of the conversational user interface. This unique flexibility serves as a proof of concept for how similar models can deeply integrate various types of data throughout different power transformer condition monitoring applications.

The rest of this paper is organized as follows. Section II explains the preliminary work done related to this research and further illustrates the efficacy of the Transformer architecture. Section III provides an overview of the proposed solution and its potential impacts on smart grid systems. Section IV details experimental results and demonstrates the proposed integrated system, including a graphical user interface. Section V concludes the research by summarizing the paper and outlining the scope of future work.

II. Background and Preliminary Work

A. LSF-Models for Smart Grids

Within the past few years, the popularity, efficiency, and universality of LSF-Models have expanded rapidly. Foundation models attempt to encompass multiple facets of data within a particular network while remaining generally applicable to multiple tasks within the same sector. Models such as Bidirectional Encoder Representation from Transformers (BERT) [6], Enhanced Representation through kNowledge IntEgration (ERNIE) [7], and Large Language Model Meta AI (LLaMA) [8] have showcased unique and successful approaches to large-scale NLP applications. Additionally, the GPT series has pushed the boundaries of the Transformer architecture through the development of increasingly complex optimization structures [9] [10].

Despite the historic success of NLP and CV LSF-models, complex smart grid systems require flexible deep learning models that can appropriately interpret several modes of data, including signals, images, videos, and text information. Even the textual mode requires the analysis of various data structures such as maintenance records, ongoing maintenance work orders, and culminating project reports. Therefore, holistic assistive systems require a deep integration of multiple machine-learning models. Current research into such a capable multimodal model is still years behind comparable designs from the NLP and CV fields [11]. The majority of proposed solutions include wavelet-based convolutional neural networks (CNNs) [12] [13] and recurrent neural networks (RNNs) [14] which both contain significant disadvantages for smart grid system applications.

B. Time-Series Analysis Techniques

The initial steps of this research included evaluating a subset of a three-phase time-series signal-based dataset [15]. The overall workflow of these algorithms was inspired by previous analysis of various machine-learning classifiers for signal classification using Discrete Wavelet Transform (DWT) decomposition for feature extraction [16]. DWT involves the convolution of a discrete signal and a shorter, pre-selected wavelet. The two signals are multiplied at increasing distances from the initial position of the discrete signal. Various families of mother wavelets can be applied in various individual patterns to find the most optimal comparison for a particular application and signal type. The convolution output passes through a series of high- and low-pass filters which produce waveforms of detail and approximation coefficients, respectively. These waveforms represent the high- and low-frequency components of the original signal. Signal anomalies are exaggerated in these decomposed forms due to their abnormal frequencies.

This research compared the features of each decomposed detail coefficient level and the last approximate coefficient level with the corresponding levels of other signals. Twelve statistical features including entropy, 5th percentile, 25th percentile, 75th percentile, 95th percentile, median, mean, standard deviation, variance, and root mean square were extracted for each decomposition level and phase. For example, if performing analysis with four decomposition levels, twelve statistical features were evaluated for five sets of DWT decomposition coefficients for three phases of data. Therefore, 180 features were extracted per data file in the randomly selected training set. More decomposition levels produced a greater number of total features which increased training time but also increased overall model accuracy. Due to diminishing returns, only five decomposition levels were used for our tests.

C. Transformer Architectures for Fault Classification

The historical success of the Transformer architecture for large text-based multiclass classification for projects such as ChatGPT prompted the search for a custom architecture capable of accurate signal-based classification. At the beginning of 2023, electrical engineering researchers at the National Institute of Technology Calicut constructed two models based on the Transformer architecture. The research compared a basic "vanilla" Transformer to an innovative Differential Architecture Search (DARTS) algorithm which produced accurate results for detecting 23 fault types and 15 fault locations for power transformers [18]. The baseline model preprocesses the input data by equalizing the samples per class and converting the arrays into TensorFlow tensors. This preprocessed data is transformed into attention embeddings through several levels of multi-head attention mechanisms and feedforward networks. These methods analyze a series of 'queries', 'keys', and 'values' to calculate the relative importance or connectivity between individual data points. This process mirrors the structure of the foundational Transformer architecture encoder. Although proven to be a more efficient method for deep data analysis than RNNs, the mathematical processes of the attention encoder are not time-dependent. In order to retain time-dependent information, the model also applies positional embeddings to the data. This allows the context of the data to be reconstructed directly following the completion of the main attention encoder process.

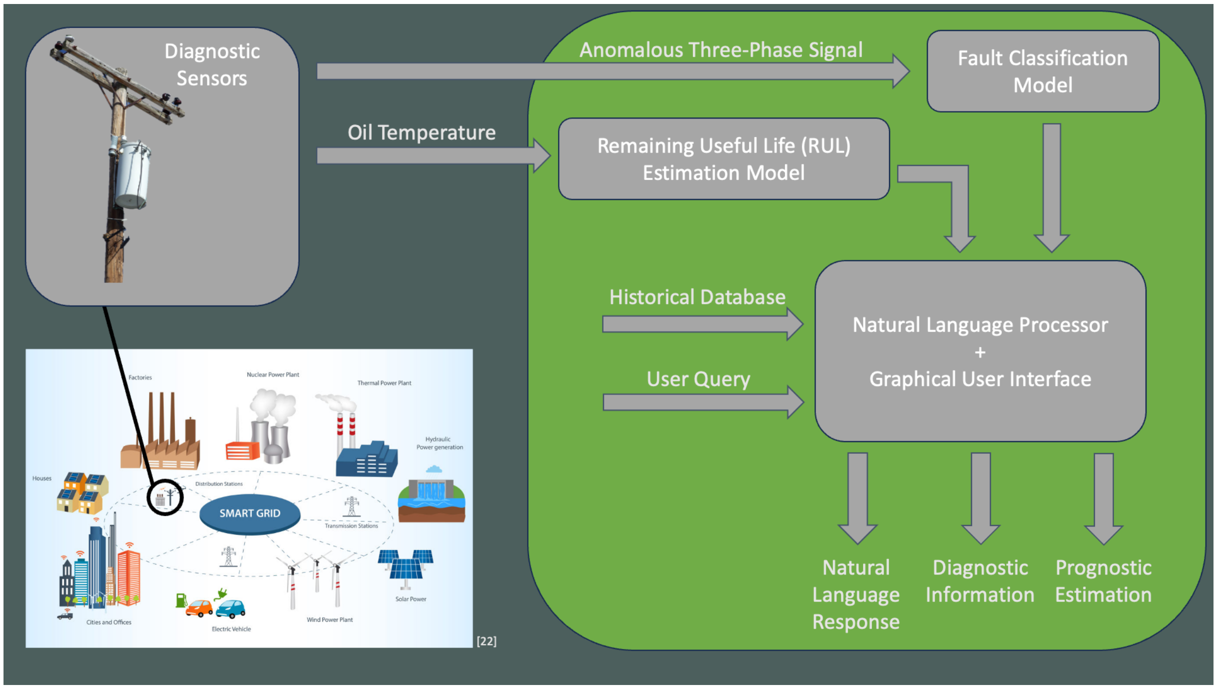

Fig. 1: Scope of proposed solution showcases multi-modal data integration. Smart grid image sourced from [19].

III. Problem Formulation and Proposed Solution

A. Research Objective

The existing approach employed by energy corporations for their prognostics and health management (PHM) system entails deploying personnel to the field to perform root cause analysis on reported issues and defective components. However, this common practice has several significant downsides, primarily in terms of its expensive and time-intensive nature. Additionally, during natural disasters and other emergencies, these inefficiencies are particularly evident and make it even more difficult to provide consumers with responsive aid in such circumstances.

This reactive maintenance technique has a notable drawback in that it is dependent on a problem being apparent before any corrective action is taken. This reactive framework often causes significant delays in the process of system-wide rehabilitation. As a result, the entire process is drawn out, and the malfunction may continue until the on-site assessment and subsequent repair attempts are exhausted. This complex situation highlights the need for energy firms to adopt a more simplified and proactive strategy in the field of PHM. Transitioning towards a technologically advanced solution, such as a self-monitoring diagnostic system that can remotely assess transformer health, would significantly mitigate several of the aforementioned challenges. By leveraging real-time data analytics and predictive algorithms, energy companies could swiftly identify irregularities or potential issues, enabling them to proactively perform maintenance or corrective measures, thus curtailing downtime and further reducing operational costs.

B. Proposed Solution

We set out to address these concerns with the goal of creating a ChatGPT-like interface that is capable of quickly identifying problems within seconds or less. The initial steps of this project included the development of a specialized fault classification function, which served as the foundation for our chatbot platform. Our initial approach leaned on conventional machine-learning techniques, encompassing the use of random forest, gradient boosting, and support vector classification machines. In order to efficiently observe abnormalities in three-phase signals, the DWT decomposition feature extraction techniques mentioned above in Section 2 were applied. Through rigorous testing, data analysis, and iterative enhancements, we steadily advanced the fault classification performance of the overall system. It was a significant achievement to see all three traditional machine-learning methods consistently deliver an admirable accuracy level of 90% on average. This represented a promising start, showcasing the potential for data-driven solutions to effectively address the PHM difficulties currently faced within the energy sector.

However, the limits of precision and efficiency restricted the machine-learning models to simple binary and small multiclass classifications. In pursuit of a more robust model that could handle tens of classes of power transformer faults and transient disturbances, we turned our attention to the Transformer architecture which had displayed historical success in several multiclass classification LSF-Models. The decision to embrace the Transformer architecture represented a strategic pivot in our approach, and the final overall proposed solution is shown in Fig. 1. This transition was marked by a transformation in our fault classification system's capabilities. By deploying the

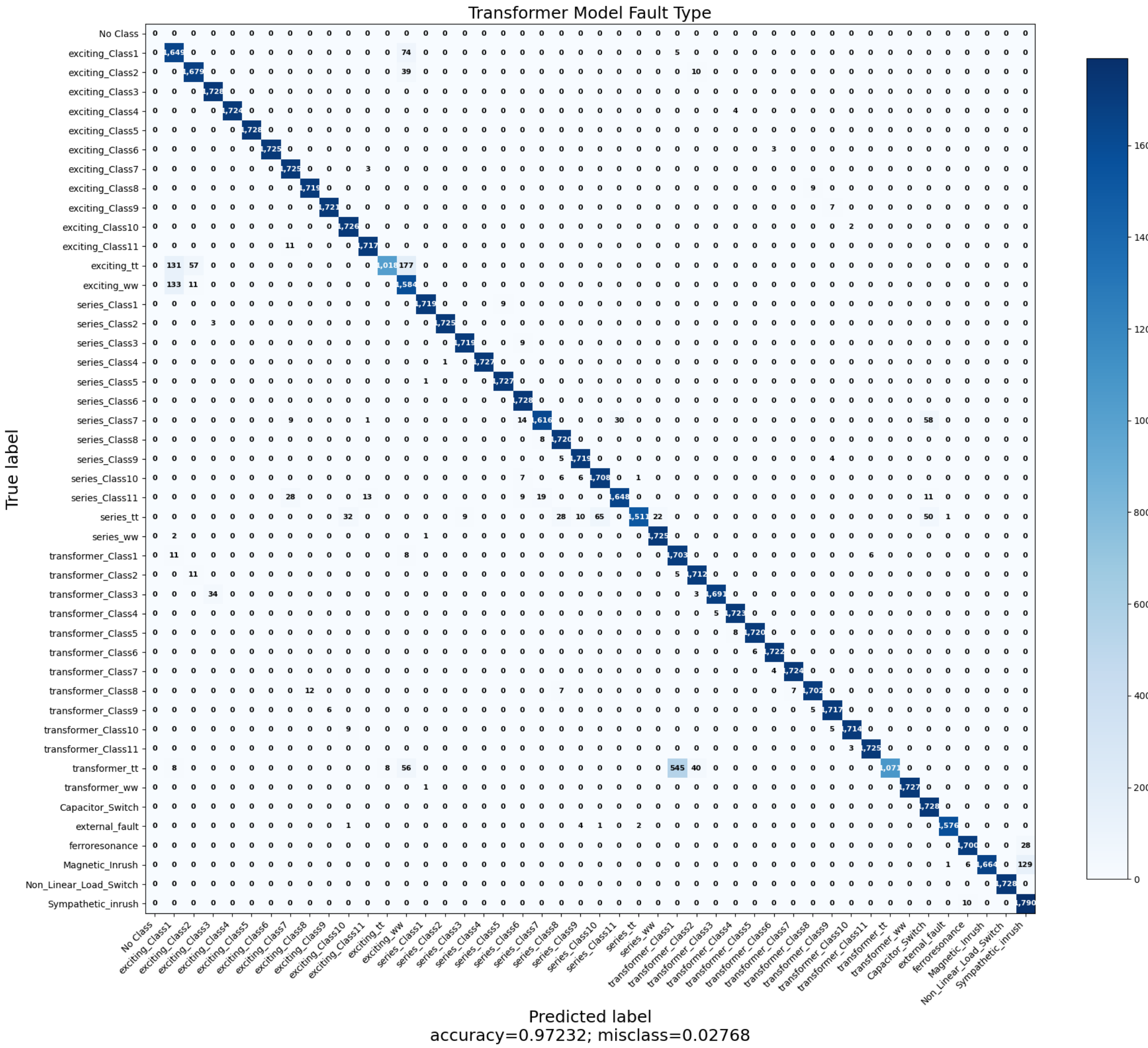

Fig. 2: The confusion matrix for our findings. Despite the model's impressive overall performance, it does have a few false positives and false negatives. These outliers could be attributed to the complex phase signals that can mimic other classes closely or vice versa.

Transformer architecture, we not only significantly improved our fault classification system's accuracy to 97.2% but also laid the foundation for a more proactive and responsive PHM framework. This important change served as evidence of the great potential for advanced PHM machine-learning architectures within the energy sector. What made this transformation even more remarkable was the adaptability of the Transformer architecture itself. Through modifications and tailored adjustments to the Transformer model we employed, we were able to fine-tune it to accurately represent our dataset. This strategic fine-tuning process allowed us to improve overall generalization results and showcased the flexibility and versatility of cutting-edge machine-learning technologies.

C. Modifications to Transformer Architecture

The Transformer architecture for signal-based multiclass classification from [18] was originally hard-coded to be used in conjunction with a proprietary private dataset. Without access to this dataset, our research utilized reverse engineering to extract the key components of the open-source code. To prove that this type of model could be used for a range of applications related to condition monitoring of power transformers, the program was dynamically generalized for a diverse range of data formats. During the initialization of the Transformer model, it was programmed to generate an appropriate number of dense and dropout layers according to the structure of the inputted dataset. These custom values were also incorporated into the input dimensions of the attention embeddings. Constant variable initializations for the number of classes and training epochs were both added to the Transformer construction and training process for future convenience. For easy integration with the user interface, functions for saving the training history, training checkpoints, and the completely trained model were also included.

IV. Results and System Demonstration

A. Transformer Model Performance

The predictive Transformer model was designed to analyze three-phase signals recorded over time for feature extraction and classification. It aimed to detect and classify faults among 45 different power transformer fault types using a dataset of over 110,000 simulated examples [15]. Our model achieved an impressive overall accuracy of 97.2% in classifying fault signals, determined through Matthew's Correlation Coefficient (MCC). As illustrated in Fig. 2, the confusion matrix exhibits strong accuracy across most classes, with a few exceptions. Notably, the outlier classes encompassed *exciting-tt*, *exciting-ww*, *transformer-Class1*, *transformer-tt*, *magnetic-inrush*, and *series-tt*. Among the *exciting* categories, significant outliers involved classifying *exciting* turn-to-turn faults and *exciting* winding-winding faults as *exciting-Class1* and *exciting-Class2*, and vice versa. Another substantial outlier occurred while classifying *magnetic inrush* as *sympathetic inrush*. The most pronounced anomaly was the misclassification of *transformer turn-to-turn* as *transformer-Class1*. Further analysis revealed that the most inaccurate predictions stemmed from the *exciting turn-to-turn* and *transformer turn-to-turn* classes, which exhibited a misclassification rate of 65%. In light of these findings, our study demonstrates the feasibility and efficiency of classifying large-scale phase signals, while recognizing the presence of certain outlier classes. Viewing the concept of predictability and fault classification through the Transformer architecture demonstrates its potential for being an accurate and reliable mechanism for multiclass classification of three-phase time-series signals found in smart grid systems.

B. Facilitating Model Interaction

To showcase the diagnostic capabilities of the Transformer architecture, we developed a Linux-based graphical user interface (GUI), shown in Fig. 3. The primary page of the GUI enables dynamic interaction with the chatbot. Key components include:

- **Input Box:** Users can input text for communication with the chatbot. The "send" button submits text to the chatbot, while the "clear screen" button begins a new conversation on a clear screen.
- **Conversation Viewer:** Displays user queries and chatbot responses in real-time.
- **Graph Viewer:** Presents graphs of classified three-phase time-series signals, aiding user comprehension and visualization.
- **History Log (not shown):** Logs conversations and corresponding graphs for future review.

C. Interface Application: Power Transformer Fault Diagnosis

Utilizing the Transformer's predictive power, users engage with the chatbot through the interface for fault diagnosis. By uploading signal files, the powerful predictive model analyzes and classifies the data. The conversational chatbot responds to user queries and instructions with the following aspects:

User greeting and request of fault diagnosis – The chatbot greets users and responds to prompts, initiating interactions. The model performs text-based classification to predict whether the user wants to converse, perform fault diagnosis, or any other programmed task. Depending on the user input, the interface appropriately replies back or initializes the requested fault classification prediction model.

User prompted to upload three-phase signal files – Users are prompted to upload an anomalous three-phase signal file from the local drive into the predictive model.

Display results – A graph of the anomalous signal is displayed above the logo, and the interface conversationally communicates the predicted fault type.

Conversational nuances – The chatbot accommodates conversational nuances and context, enabling user-friendly interactions that account for spelling mistakes or grammatical errors. Note how the interface correctly classifies the request for fault diagnosis while "fault diagnosis" is misspelled.

V. Conclusions and Discussions

The presented study introduces a predictive Transformer model designed to analyze three-phase signals for power transformer fault classification. The model's commendable accuracy of 97.2% in identifying fault types underscores its efficacy. Leveraging MCC, we are poised to discuss finer metrics that delve deeper into the model's performance, which aligns with an improvement from our initial machine-learning models. Our results are promising, with the majority of the confusion matrix aligning accurately with expectations. The high overall accuracy, although impressive, is tempered by the challenges posed by outlier classes. Notably, *transformer* and *exciting* classes exhibit inconsistencies in classification that require further investigation into their outliers and erroneous patterns. Although misclassifications exist, they serve as valuable insights into areas where the model's sensitivity can be refined. This study emphasizes that while automated fault classification is obtainable, continuous refinement is crucial to address nuanced variations in fault behavior.

The development of a Linux-based GUI further enhances the accessibility of the Transformer model. The interface offers a user-friendly platform for dynamic interaction and fault diagnosis, amplifying the model's applicability to real-world scenarios. In the future, other enhancements to our interface could prove crucial to improving our predictive model, such as expanding the framework to include real-time online sensor condition monitoring in addition to cross-

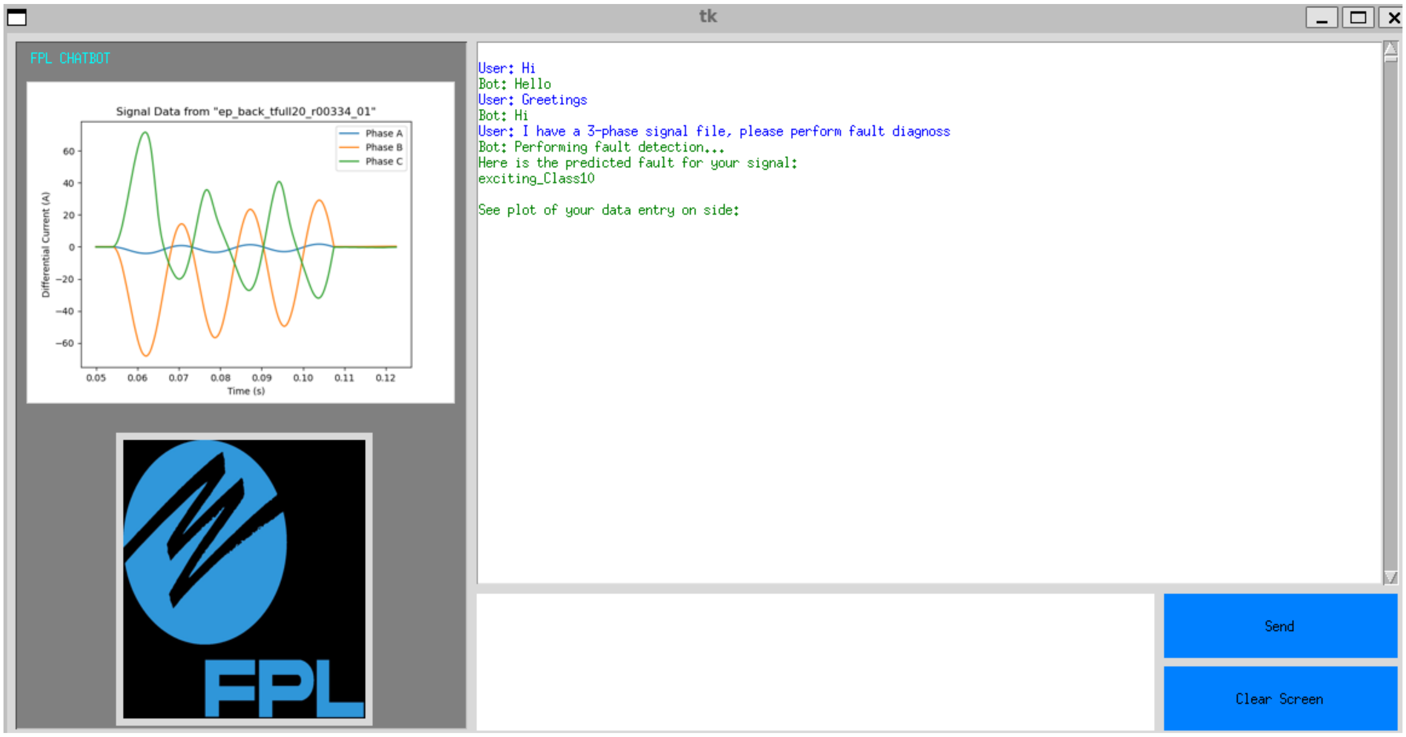

Fig. 3: Our GUI after a fault diagnosis. The predictive model uploads the graphed signals shown on the left and produces the fault class in the conversation window on the right. Users have the option to clear the screen or continue responding in the input box below. As shown above, the conversational nature of the GUI understands context and even with spelling errors, fault diagnostics can still be activated.

referencing with historical prognostic databases. In conclusion, our work contributes to the evolving landscape of PHM and fault diagnosis for smart grid power transformers. This research serves as a proof-of-concept solution, showcasing how a single user interface can potentially improve the reliability and performance of power transformers within smart grid systems.

ACKNOWLEDGMENT

The authors would like to thank Siyuan Du for his contributions to the GUI and general research assistance.

REFERENCES

[1] U.S. Energy Information Administration - EIA - independent statistics and analysis. Use of electricity - U.S. Energy Information Administration (EIA). (2023, April). https://www.eia.gov/energyexplained/electricity/use-of-electricity.php

[2] 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.

[3] L. Ouyang, J. Wu, X. Jiang, D. Almeida and C. Wainwright, et al., "Training language models to follow instructions with human feedback," in Proc. Advances in Neural Information Processing Systems, 2022, pp. 27730-27744.

[4] A. Kirillov, E. Mintun, N. Ravi, H. Mao and C. Rolland, et al., "Segment Anything," [Online]. Available: https://arxiv.org/abs/2304.02643.

[5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit and L. Jones, et al., "Attention is All You Need," Advances in neural information processing systems, vol. 30, 2017.

[6] J. Devlin, M. Chang, K. Lee and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," [Online]. Available: https://arxiv.org/abs/1810.04805.

[7] Y. Sun, S. Wang, Y. Li, S. Feng and X. Chen, et al., "ERNIE: Enhanced Representation through Knowledge Integration," [Online]. Available: https://arxiv.org/abs/1904.09223.

[8] H. Touvron, T. Lavril, G. Izacard, X. Martinet and M. Lachaux, et al., "Llama: Open and efficient foundation language models," [Online]. Available: https://arxiv.org/abs/2302.13971.

[9] A. Radford, K. Narasimhan, T. Salimans and I. Sutskever, "Improving language understanding by generative pre-training," [Online]. Available: https://openai.com/blog/chatgpt.

[10] OpenAI, "GPT-4 Technical Report," [Online]. Available: https://arxiv.org/abs/2303.08774.

[11] Li, Y., Wang, H., and Sun, M. (2023). ChatGPT-Like Large-Scale Foundation Models for Prognostics and Health Management: A Survey and Roadmaps. [Online]. Available: https://arxiv.org/abs/2305.06472

[12] H. Wang, Z. Liu, D. Peng and M.J. Zuo, "Interpretable convolutional neural network with multilayer wavelet for Noise-Robust Machinery fault diagnosis," Mech. Syst. Signal Pr., vol. 195, pp. 110314, 2023.

[13] Y. Wang and H. Wang, "Wavelet attention-powered neural network framework with hierarchical dynamic frequency learning for lithium-ion battery state of health prediction," Journal of Energy Storage, vol. 61, pp.106697, 2023.

[14] H. Liu, J. Zhou, Y. Zheng, W. Jiang and Y. Zhang, "Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders," ISA T., vol. 77, pp. 167-178, 2018.

[15] Pallav K Bera, Can Isik, Vajendra Kumar, February 13, 2020, "Transients and Faults in Power Transformers and Phase Angle Regulators (DATASET)", IEEE Dataport, doi: https://dx.doi.org/10.21227/1d1w-q940.

[16] "A guide for using The wavelet transform in machine learning," ML Fundamentals, https://ataspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning/ [Accessed Aug. 29, 2023].

[17] F. Wasilewski, "Discrete wavelet transform (DWT)," Discrete

Wavelet Transform (DWT) - PyWavelets Documentation, https://pywavelets.readthedocs.io/en/latest/ref/dwt-discrete-wavelet-transform.html [Accessed Aug. 29, 2023].

[18] J. B. Thomas and S. K.V., "Neural architecture search algorithm to optimize deep transformer model for fault detection in Electrical Power Distribution Systems," Engineering Applications of Artificial Intelligence, vol. 120, p. 105890, 2023. doi:10.1016/j.engappai.2023.105890

[19] "The Smart Grid could hold the keys to electric vehicles," IEEE Innovation at Work, https://innovationatwork.ieee.org/the-smart-grid-could-hold-the-keys-to-electric-vehicles/ [Accessed Aug. 29, 2023].