Energy Proceedings
Vol 52, 2025
ISSN 2004-2965

Intelligent Fault Diagnosis for Overhead Lines with Covered Conductors: Using Large Language Model

Genghong Lu1, Siqi Bu 2*
1 Department of Electrical and Electronic Engineering, Centre for Advances in Reliability and Safety, The Hong Kong Polytechnic University
2 Department of Electrical and Electronic Engineering, Shenzhen Research Institute, Research Centre for Grid Modernisation, International Centre of Urban Energy Nexus, Centre for Advances in Reliability and Safety, Research Institute for Smart Energy, and Policy Research Centre for Innovation and Technology, The Hong Kong Polytechnic University
(Corresponding Author: siqi.bu@polyu.edu.hk)

ABSTRACT

Fault diagnosis of partial discharge (PD) is crucial for the protection of overhead lines with covered conductors. Facing the challenge of identifying PDs that may have diverse fault patterns from background noise interferences, a novel intelligent fault diagnosis utilizing the large language model (LLM) is developed. To effectively apply LLM to PD diagnosis, the domain knowledge-based prompts are designed by incorporating the specific domain information, PD detection task description, and measurement data information. To further improve the capability of LLM reasoning antenna signals, a signal reprogramming method is adopted to align the modalities of the measured signals and natural language. Finally, an output projection is constructed to identify PD by taking in the features learned from the LLM, whose backbone model remains intact during the learning process. Experimental results validate the efficiency and effectiveness of the developed method.

Keywords: intelligent fault diagnostics, large language model, partial discharges, power line protection.

1. INTRODUCTION

Fault diagnosis of power lines plays a critical role in improving the reliability of power distribution systems [1]. Faults caused by contact of covered conductors (CCs) with tree branches or CC falling onto the ground often result in insulation deterioration of CC [2], [3]. A key indicator for insulation deterioration of CC is the partial discharge (PD) activity [4]. PDs are small electric sparks or discharges caused by an electric field enhancement [5]. Unlike the signals collected in the laboratory, the raw signals from real overhead lines include uncertain information caused by external background noise interference. One of the major challenges for PD pattern recognition is to extract PD signals from a large number of background noises [6].

Machine learning methods, such as random forest [6], and light gradient boosting machine (lightGBM) [7], have been used for detecting PD signals. To identify fault peaks in fault PD patterns, fault indicators are designed, and the corresponding statistical features are calculated. With the customized features, machine learning methods are used as classifiers to recognize PD patterns. To characterize PD-related pulse shapes, the clustering approach is used as a feature extractor; then, lightGBM is implemented as a classifier for CC fault detection [7]. Considering that manual feature preparation is time-consuming, intelligent fault diagnostics based on deep learning methods have been developed [8]-[10]. An ensemble deep learning framework combining base learners and a meta-learner is proposed to automatically learn features from antenna signals and classify PD signals[11].

Although DL methods provide promising solutions to learn and recognize fault patterns from raw signals, they require a high level of knowledge of both the DL algorithms and the application industry. In addition, the application of DL models to various fault diagnosis tasks necessitates the training of numerous models, which is time-consuming and laborious. Therefore, intelligent fault diagnosis utilizing a simple training and application strategy can enhance user-friendliness and mitigate the challenges associated with implementing DLs in practical applications.

Large language models (LLMs) are powerful tools with robust pattern recognition and reasoning abilities. Prompt-based learning, which uses frozen LLMs to tune trainable prompt embeddings, provides an efficient

This is a paper for the 16th International Conference on Applied Energy (ICAE2024), Sep. 1-5, 2024, Niigata, Japan.

solution to employ LLMs in different downstream tasks [12]. Therefore, the intelligent fault diagnosis utilizing prompt-based learning with LLMs necessitates only the fine-tuning of embedding layers and output projection layers, offering a streamlined approach for implementing LLMs across diverse fault diagnosis scenarios.

In addition to the implementation of LLMs in computer vision and natural language processing, the capabilities of LLMs in the electric energy sector[13], such as load forecasts and power flow data analysis, have been explored. By leveraging the natural language understanding ability and knowledge reasoning ability of LLM, an LLM-based fault identification and troubleshooting solution is developed [14]. The superior pattern recognition and reasoning capability make LLMs promising solutions for intelligent Fault diagnosis. However, the inherent difference between measurements (e.g., antenna signals) and natural language poses significant challenges for LLMs in directly identifying fault patterns from measurements, ultimately hindering their practical industry application.

To address the above challenge, a novel LLM-based intelligent fault diagnosis for overhead lines with CCs is developed. Domain knowledge-based prompts, which consist of distinct domain characteristics, specific learning tasks, and enriched data information, are designed to be used as a main part of inputs. The measured signals are reprogrammed and used as another part of inputs. This method predicts faulted signals by using the frozen LLM and tuning the trainable parameters in signal reprogramming layers and output projection layers. The contributions of this paper are as follows.

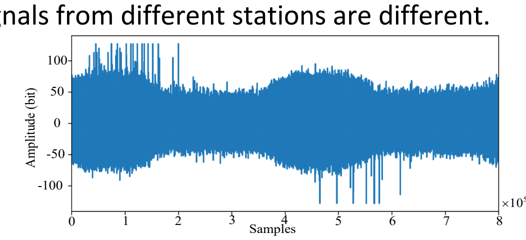
1. Compared to the development of DL-based fault detection methods requiring a high level of knowledge of both the DL algorithms and the application industry, a novel intelligent fault diagnosis adopting a simple application strategy, which uses frozen LLM to tune a small number of trainable parameters, is developed. By exploring the PD detection performance of LLM, this study shows the potential of employing LLMs in addressing complex industry application challenges.
2. Facing the inherent difference between measurements (e.g., antenna signals) and natural language, which would degrade the performance of LLM reasoning measured signals, a signal reprogramming method is adopted to align the modalities of the measured signals and natural language. In addition, domain knowledge-based prompts are designed as a task-specific activation of the LLM to enhance the inference capability of LLM.

2. FAULT DIAGNOSIS FOR OVERHEAD LINES WITH COVERED CONDUCTORS

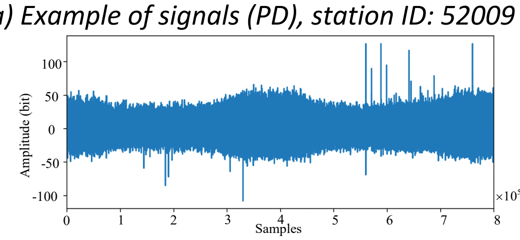
2.1 *CC fault diagnosis problem*

CC line contacts with surrounding vegetation would cause faults, such as phase-to-ground and phase-to-phase faults [2]. These faults are accompanied by PD activities caused by inhomogeneous electrical fields. The occurrence of PD activities would cause the degradation of CC, leading to catastrophic results, e.g., power supply cuts and forest fires. Therefore, PD diagnosis plays a critical role in securing the power distribution systems.

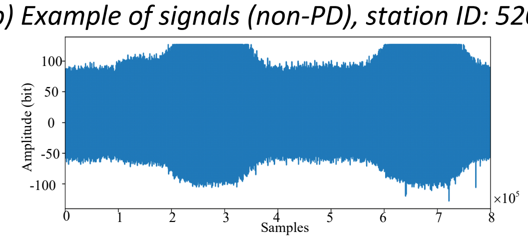
To detect PD activities for CC fault diagnosis and monitor the condition of CC insulation systems, a contactless antenna method is used [15]. Unlike the data from the laboratory, the data collected from different stations in the real environment are very noisy. Examples of antenna signals collected using the contactless antenna method are shown in Fig.1. By observing the non-PD signals collected from stations 52009 (Fig.1 (b)) and 52010 (Fig.1 (d)), it can be found that the magnitudes of signals from different stations are different.



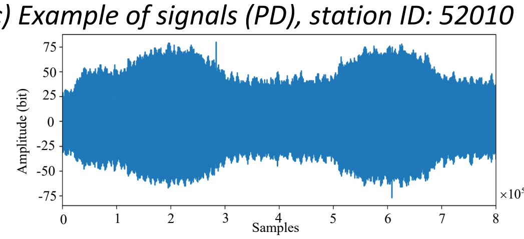
(a) Example of signals (PD), station ID: 52009



(b) Example of signals (non-PD), station ID: 52009



(c) Example of signals (PD), station ID: 52010



(d) Example of signals (non-PD), station ID: 52010

Fig. 1 Examples of signals from contactless sensors.

Furthermore, it can be observed that the signals contain high levels of noise, including random pulse interference (RPI) and discrete spectral interference (DSI)[16].

DSIs are the narrowband noise spectrums that primarily originate from wireless communication systems, including communication and radio systems. When the DSIs occur, PD signals are modulated on DSI signals, decreasing the ratio of the PD pattern. Additionally, the amplitudes of DSIs fluctuate in response to external factors, such as weather conditions. A high level of DSI noise can cover the PD pattern and complicate the classification task.

RPI noises are caused by any CC non-relative pulse appearance in the examined signal. There are various sources of this interference including lightning strikes and switching operations. Distinguishing between the signal patterns of non-PD signals and actual PD signals proves challenging due to the resemblance in pulse frequency characteristics of false hit peaks generated by RPI and PD signals, often leading to misinterpretation.

In addition to noise interference, the diversity of PD patterns, see Fig. 1 (a) and (c), makes the fault diagnosis a more challenging problem. The PD-pattern is the time pattern of the PD activity, influenced by a range of external factors like the shape and length of tree branches, the number of contact points between tree branches and CC, and changing environmental conditions like temperature, among others. Consequently, the complexity of establishing the correlation between raw data and partial discharge presents a significant challenge in developing an intelligent fault diagnosis.

2.2 Overview of the proposed intelligent fault diagnosis

The learning problem of the LLM-based fault diagnosis using LLM can be expressed as

$$y = LLM(\boldsymbol{x}), \tag{1}$$

where LLM indicates the pre-trained model, such as generative pre-trained transformer (GPT), GPT2, and bidirectional encoder representation from transformers (BERT). $\boldsymbol{x}$ indicates the signals and y is the binary detection results. If the fault is detected, $y = 1$. Otherwise, $y = 0$.

However, due to the lack of specific domain knowledge associated with PD detection in CC lines, the detection performance of LLM may not be satisfactory. Therefore, the domain knowledge $\boldsymbol{k}$ is extracted to be input to the LLM. The developed LLM-based intelligent fault diagnosis framework for CC lines is illustrated in Fig. 2.

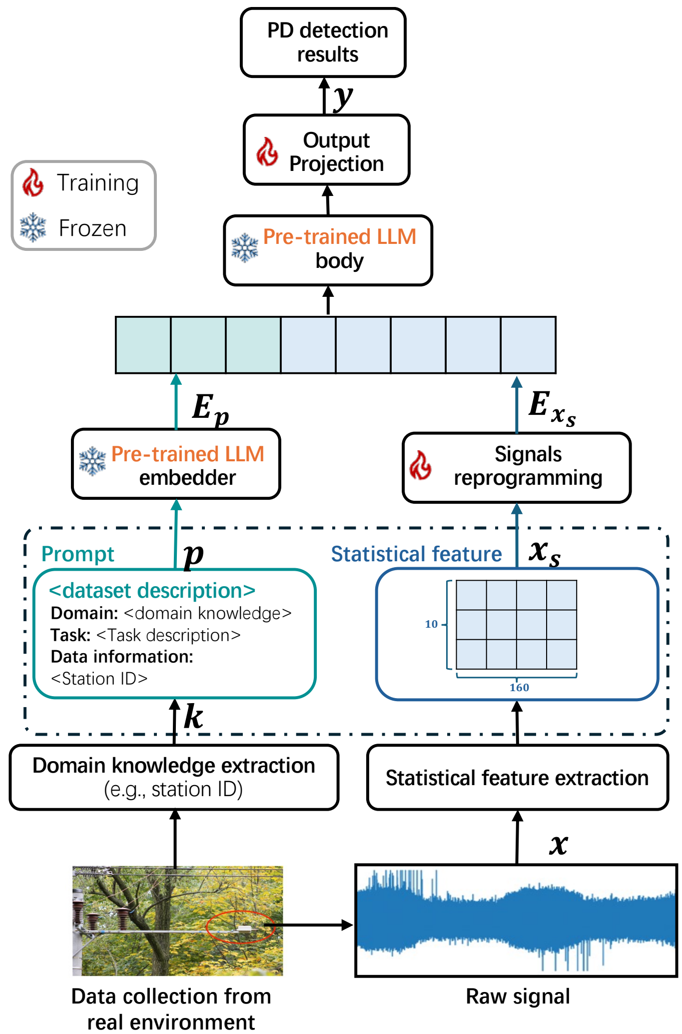


Fig. 2 The overall framework of intelligent fault diagnosis using LLM

Data is first collected from the real environment using the antenna contactless method. The domain knowledge $\boldsymbol{k}$, including background associated with PD detection in CC lines and data information, is extracted and constructed as prompts $\boldsymbol{p}$. The pre-trained LLM embedder is used to embed the prompts $\boldsymbol{p}$. In the meanwhile, the collected raw signals $\boldsymbol{x}$ are preprocessed to remove noise and extract statistical features $\boldsymbol{x_s}$. The signals reprogramming is adopted to process the extracted features $\boldsymbol{x_s}$. The frozen LLM takes the embedded prompts $\boldsymbol{E_p}$ and features $\boldsymbol{E_{x_s}}$ as inputs and outputs the hidden representations to the output projection, which generates the final detection results $\boldsymbol{y}$.

Therefore, the learning problem of the developed intelligent fault diagnosis using LLM is described as

$$y = LLM_\theta(\boldsymbol{k}, \boldsymbol{x}), \tag{2}$$

where θ indicates the trainable parameters in signal reprogramming and output projection layers.

3. PROPOSED METHODOLOGY

3.1 Domain knowledge-based prompt embedding

Prompting functions as an effective task-oriented activation of LLMs [17]. Prompt-as-Prefix can enrich the input context and enhance the LLM's adaptability to downstream tasks. Therefore, the domain knowledge of measurement and expert-knowledge-based PD detection results are used for constructing prompts.

Fig. 2 gives an example of prompts, which consists of three components: domain information, task description, and data information. Domain information provides LLMs with background relevant to PD detection. Task description provides the LLM with task instructions, e.g., fault detection given antenna signals. Since antenna signals from different stations are different, the station ID is used as one of the data information to enrich the input information. In this way, the prompts incorporated with distinct domain characteristics, specific learning tasks, and enriched data information, are designed to enhance the fault pattern recognition capabilities of LLMs.

Then, the embedder of the pre-trained LLM is used to embed the prompts $\boldsymbol{p}$.

$$\boldsymbol{E_p} = Embed_{LLM}(\boldsymbol{p}) \tag{3}$$

where $Embed_{LLM}$ indicates the process of tokenization and token embedding of the LLM. For different LLMs, e.g., GPT2 and BERT, the $Embed_{LLM}$ can be different.

3.2 Raw signals processing and reprogramming

3.2.1 Statistical feature extraction

Since the antenna signals contain a high amount of background noises, a denoising procedure using univariate wavelet denoising (mother wavelet: db4, level of decomposition: 1) and hard thresholding is applied to suppress the noises. The signal is decomposed into first level of detail and approximation coefficients. Then, the hard thresholding is performed on the detailed coefficients (c_d) to remove noises. The threshold is calculated as:

$$\lambda = \frac{MAD(|c_d|)}{0.6745\sqrt{2\log(N)}}, \tag{4}$$

where N is the number of samples, MAD refers to the mean absolute deviation. The coefficient values greater than λ will be retained; otherwise, the co-efficient values will be set to 0.

The denoised signals are divided into 160 windows with 5000 samples each. In each window, 10 statistical features are extracted, including the mean values, standard deviation, percentile values, and a percentile rank of mean. When calculating the percentile values, the i^{th}, ($i = 0,1,25,50,75,99,100$) percentile values of the signals are considered as the features. For example, when $i = 0$, the minimum value of the signal will be extracted as a feature. Therefore, statistical features $\boldsymbol{x_s}$ (i.e., a matrix of 160×10) will be extracted from each signal and used as inputs of the PITCN. These features represent statistical information of antenna signals.

3.2.2 Signals embedding

To align the modalities of antenna signals and natural language, the antenna signals are reprogrammed into text prototype representations, thereby enhancing their compatibility with the linguistic capabilities of language models.

A multi-head cross-attention layer is adopted to reprogram statistical features $\boldsymbol{x_s}$.

$$\boldsymbol{E_{x_s}} = SoftMax\left(\frac{\boldsymbol{Q_k^i K_k^i}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V_k^i} \tag{5}$$

where $\boldsymbol{Q_k^i} = \boldsymbol{x_s w_k^Q}$ is query matrices, $\boldsymbol{K_k^i} = \boldsymbol{E w_k^K}$ is key matrices, and $\boldsymbol{V_k^i} = \boldsymbol{E w_k^V}$ is value matrices. $\boldsymbol{E}$ denotes the pre-trained word embedding, $\boldsymbol{w_k^Q}$, $\boldsymbol{w_k^K}$, and $\boldsymbol{w_k^V}$ are trainable parameters. d_k is the dimension of the head.

3.3 LLM processing

After concatenating the embedded prompts and statistical features, i.e., $\boldsymbol{E_{in}} = [\boldsymbol{E_p}; \boldsymbol{E_{x_s}}]$, the pre-trained LLM is used to output the learned representations.

$$\boldsymbol{x_{LLM}} = LLM(\boldsymbol{E_{in}}) \tag{6}$$

In this paper, the pre-trained BERT is utilized. The model architecture of BERT is a multi-layer bidirectional transformer encoder [18]. To be more specific, the $\text{BERT}_{\text{BASE}}$ with 12 transformer blocks, and 12 self-attention heads is used. Its hidden size is 768 and the total parameters are 110M. Note that the embedder and body of LLM are frozen during the fine-tuning.

3.4 Output generation

Finally, the learned representations $\boldsymbol{x_{LLM}}$ are flattened and sent to an output projection, which consists of four fully connected layers.

$$\boldsymbol{x^l} = \sigma(\boldsymbol{w^l x^{l-1}} + \boldsymbol{b^l}), l = 1, \dots, L-1 \tag{7}$$

$$\boldsymbol{Y_{out}} = \boldsymbol{w^L x^{L-1}} + \boldsymbol{b^L} \tag{8}$$

where $\boldsymbol{w^l}$ and $\boldsymbol{b^l}$ are trainable parameters of l^{th} layer. σ is the non-linear activation function, i.e., ReLU.

4. CASE STUDIES

4.1 Data description

The signals collected via a contactless antenna method in CC medium-voltage overhead power transmission lines are used for validation [15]. The examples of raw signals are shown in Fig.1. The total 1700 data (180 series containing PD signals and 1520 series without PDs) are used for training and testing. Each series consisted of 800,000 observations. The magnitude of signals is in the range from −128 to 127. 70% of the data is used for training and the rest is used for testing.

4.2 Model performance comparison

To validate the PD detection efficiency of the developed LLM-based method, the dataset is used for testing model performance of the expert-knowledge-based PD detection[16], XGBoost [19], ensemble deep learning [11], and the developed method. The confusion matrices are shown in Fig. 4.

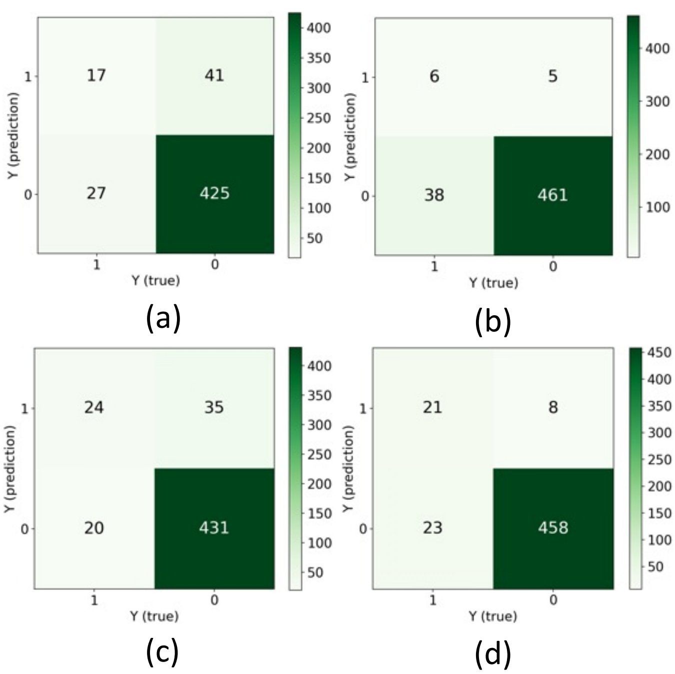


Fig. 4 Confusion matrices using different methods. (a) Expert-knowledge-based PD detection. (b) XGBoost. (c) Ensemble DL. (d) Proposed LLM-based method

In the confusion matrices, class 1 (positive) indicates fault, and class 0 (negative) indicates normal signals. The test dataset consists of 44 faulted samples and 466 normal samples. To evaluate PD detection performance on the imbalanced dataset, the Matthews correlation coefficient (MCC) is used as the main evaluation metric:

$$MCC = \frac{(TP \times TN - FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (9)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. TP, TN, FP, and FN can be obtained by calculating the confusion matrix. To give a more comprehensive comparison of detection performance, other metrics, such as precision, recall, F1 score, and accuracy are given in TABLE I.

TABLE I Detection performance comparison among different methods

Methods	Precision	Recall	F1	MCC	Accuracy
expert-knowledge-based PD detection	0.29	0.39	0.33	0.26	0.87
XGBoost	0.55	0.14	0.22	0.24	0.92
ensemble DL	0.41	**0.55**	**0.90**	0.41	0.89
LLM-based method	**0.72**	0.48	0.58	**0.56**	**0.94**

By comparing confusion matrices and metrics of different methods, it can be observed that the DL-based methods, i.e., ensemble DL and LLM-based method outperforms the other two methods. This demonstrates the efficiency of DL methods in learning and recognizing fault patterns. The ensemble DL achieves the highest recall and F1 score among the four methods. This indicates that the ensemble DL is more sensitive to positive data compared to the other methods. The developed LLM-based method has the highest precision, MCC, and accuracy. Since MCC takes into account all four values in the confusion matrix, the highest MCC indicates that the developed LLM-based method is able to predict both classes well.

5. LIMITATIONS AND FUTURE DIRECTIONS

The developed LLM-based method shows potential for enhancing the efficiency and effectiveness of fault diagnosis. Nevertheless, there are inherent limitations that require attention, as well as opportunities for further development and improvement.

One of the significant challenges in applying LLMs within fault diagnosis is the lack of domain-specific data in the pretraining of LLMs. Considering the diverse fault characteristics across various applications, it is difficult for an LLM that trained on publicly available datasets to effectively learn distinct fault patterns and solve task-specific fault diagnosis problems. In addition, the lack of interpretability poses challenges for operators in comprehending the predictions generated by LLMs.

Furthermore, the poor quality of real-environment data, such as data missing and noises, can significantly impair the inference capabilities of LLMs.

Therefore, future research should prioritize prompt engineering techniques to incorporate domain-specific knowledge into LLMs, thereby enhancing their pattern recognition and reasoning capabilities for addressing diverse fault diagnosis problems. Explainability techniques should be explored to improve the interpretability of LLMs. To deal with real-world data problems, data augmentation techniques should be studied to generate high-quality data for fine-tuning LLMs to improve the detection performance of LLM-based fault diagnosis.

6. CONCLUSIONS

In the face of the problem of detecting diverse PD patterns from noise interferences, a novel intelligent fault diagnosis using LLM is developed. The domain knowledge associated with PD diagnosis for overhead lines with CC and the collected raw signals are processed and embedded as the inputs of LLM. The developed method adopts a simple training strategy, which uses frozen LLM to tune a small number of trainable parameters in signal reprogramming layers and output projection layers.

The experimental results show that the LLM is a promising tool for recognizing diverse PD fault patterns. The highest MCC achieved by the developed LLM-based method proves that the pre-trained language model can be effectively applied to solve industry problems by employing a simple implementation strategy.

ACKNOWLEDGEMENT

This work is supported by the Centre for Advances in Reliability and Safety (CAiRS) admitted under AIR@InnoHK Research Cluster.

REFERENCE

[1] Bo Z.Q. et al., "Developments of power system protection and control," Prot. Contr. Modern Power Syst., vol. 1, no. 1, pp. 1-8, Jul. 2016.

[2] Rosle N. et al., “Partial discharges classification methods in xlpe cable: A review,” IEEE Access, vol. 9, pp. 133258–133273, 2021.

[3] Misak S. et al., “Problems associated with covered conductor fault detection,” in 11th Int. Conf. Elect. Power Qual. Utilisation, Lisbon, Portugal, 2011, pp. 1-5.

[4] Eigner A. and Rethmeier K.,“An overview on the current status of partial discharge measurements on ac high voltage cable accessories,” IEEE Elect. Insul. Mag., vol. 32, no. 2, pp. 48-55, March-April 2016.

[5] IEEE Guide for Partial Discharge Testing of Shielded Power Cable Systems in a Field Environment, IEEE Standards, Piscataway, NJ, USA, Feb. 2007.

[6] Misák S. et al., "A complex classification approach of partial discharges from covered conductors in real environment," IEEE Trans. Dielectrics Elect. Insul., vol. 24, no. 2, pp. 1097-1104, Apr. 2017

[7] Chen K. et al., "Fault Detection for Covered Conductors With High-Frequency Voltage Signals: From Local Patterns to Global Features," IEEE Trans Smart Grid, vol. 12, no. 2, pp. 1602-1614, Mar. 2021.

[8] Gao J., Guo M., and Chen D.Y., "Fault line detection using waveform fusion and one-dimensional convolutional neural network in resonant grounding distribution systems," CSEE J. Power Energy Syst., vol. 7, no. 2, pp. 250-260, Mar. 2021.

[9] Hu J. et al., "Fault location and classification for distribution systems based on deep graph learning methods," J. Modern Power Syst. Clean Energy, vol. 11, no. 1, pp. 35-51, Jan. 2023.

[10] Liao W. et al., "Fault diagnosis of power transformers using graph convolutional network," CSEE J. Power Energy Syst., vol. 7, no. 2, pp. 241-249, Mar. 2021.

[11] Klein L. et al. “Antenna con-tactless partial discharges detection in covered conductors using ensemble stacking neural networks,” Expert Syst. Appl., vol.213, p.118910, 2023.

[12] Ni, S. and Kao, H. Y.. Masked Siamese prompt tuning for few-shot natural language understanding. IEEE Trans Artif. Intell., vol.*5*, no.2, pp.624-633, 2023.

[13] Majumder, S., et al. Exploring the capabilities and limitations of large language models in the electric energy sector. *Joule*, vol.8, no.6, pp.1544-1549, 2024.

[14] Liu, P., et al. Joint knowledge graph and large language model for fault diagnosis and its application in aviation assembly. IEEE Trans. Ind. Inform., 2024.

[15] Klein L. et al.,“A data set of signals from an antenna for detection of partial discharges in overhead insulated power line,” Scientific Data, vol.10, p.544, Aug. 2023.

[16] Fulnecek J. and Misak S., “A simple method for tree fall detection on medium voltage overhead lines with covered conductors,” IEEE Trans. Power Del., vol. 36, no. 3, pp. 1411-1417, Jun. 2021.

[17] Jin, M., et al. Time-llm: Time series forecasting by reprogramming large language models. *arXiv preprint arXiv:2310.01728*, 2023.

[18] Vaswani, A., et al. Attention is all you need. Advances neural inf. Process. Syst., 2017, 30.

[19] Martinovič T. and Fulneček J., "Fast Algorithm for Contactless Partial Discharge Detection on Remote Gateway Device," IEEE Trans. Power Del., vol. 37, no. 3, pp. 2122-2130, Jun. 2022.