

Artificial Intelligence Based Fault Diagnosis and Relay Protection Technology in Power Systems

Jingnuo Wu^{1,a}, Yao Xiao^{1,b,*}

¹State Grid Tong Hua Power Supply Company, State Grid Jilin Electric Power Company Limited,
Tonghua, Jilin, China

^a735537541@qq.com, ^b676412762@qq.com

*Corresponding author

Keywords: Artificial Intelligence, Power System Fault Diagnosis, Relay Protection Technology, Real-Time Prediction

Abstract: Nowadays, existing fault diagnosis technologies have problems such as slow response speed, low accuracy, and weak adaptive ability. To prevent overfitting, this article can use a strictly separated set of training and testing samples to train the model. In order to ensure the generalization performance of the model, mutual confirmation technology was adopted. The computing power of GPUs can be utilized to effectively process massive amounts of data and improve training efficiency. In the field of fault diagnosis, the proposed method can achieve real-time collection of the operating status of the power grid, and use the established artificial intelligence model to analyze it, thereby achieving rapid identification and localization of system fault types and locations. This method has self-learning function, which can continuously improve the accuracy of fault diagnosis while accumulating data. At the same time, the algorithm also has an alarm function, which can predict and warn the system before it malfunctions, thereby taking corresponding preventive measures. At a transmission speed of 10 kbps, the error detection accuracy of the system reached 98.5%. This article can promote the development of power grid fault diagnosis and protection technology, which is conducive to providing new ideas and methods for power system fault diagnosis and relay protection.

1. Introduction

Nowadays, with the increasing complexity of power grid structures, fault diagnosis and relay protection have become urgent tasks. Although traditional methods have achieved some results, their ability to handle diverse faults in large and highly complex power systems is not strong enough. In recent years, artificial intelligence has become a new method for power system fault diagnosis and relay protection due to its excellent data processing and recognition capabilities. The above research indicates that artificial intelligence has great application prospects in improving the efficiency of power grid operation, especially in complex dynamic systems.

This project aims to combine artificial intelligence theories and methods such as deep learning, machine learning, and data mining to study a new type of fault diagnosis and relay protection method for power systems. Firstly, this project aims to utilize big data and historical data to learn

and optimize deep neural network models, in order to improve their prediction accuracy and responsiveness. On this basis, combined with actual operating data, the article explores power system protection strategies based on intelligent algorithms to minimize the safe operation of the power grid. The expected research results are of great significance for improving the safe and stable operation of the power grid, and can provide new ideas for fault handling of similar complex industrial systems.

This article consists of three main parts. Firstly, this article conducted in-depth research on it from three aspects: data collection, model training, and simulation experiments, and analyzed it. Secondly, this project aims to further study the integration of this method with the current power grid protection system, in order to achieve real-time information processing and automation of intelligent protection strategies for the power grid. Finally, this project can conduct empirical analysis and evaluation of the proposed theories and methods through field experiments and typical case analysis, providing theoretical basis and technical support for the safe and stable operation of the power grid. This project focuses not only on technological innovation and application, but also on evaluating the potential for industrial development.

2. Related Work

It is necessary to conduct research on fault diagnosis and relay protection technology to ensure the safe and stable operation of the power grid. Current research mostly focuses on improving diagnostic speed and accuracy, such as using machine learning and other methods to analyze fault signals. Ren Guanyu discussed the fault diagnosis and handling methods of the relay protection system in power plants [1]. Xu Fei analyzed and studied the hidden faults of relay protection in the power grid [2]. Yang Yifan explored the diagnosis and on-site treatment strategies for common faults in power plant relay protection [3]. Zhu Xu studied the online monitoring and fault diagnosis technology analysis of the secondary circuit of intelligent substation relay protection [4]. Zhang Ping studied the fault diagnosis technology for power system tripping based on peephole structure LSTM [5]. However, such studies often rely on a large amount of labeled data, and when the data quality is not high or the data volume is insufficient, the generalization ability and accuracy of the model can be limited.

With the rapid development of AI technology, researchers are increasingly concerned about its application in power system relay protection. Especially deep learning techniques are seen as tools that can process complex power grid data and achieve efficient fault prediction. Wang Da explored the design and application scheme of a secondary fault intelligent diagnosis system based on regulatory cloud data analysis [6]. Wang Kun conducted fault analysis and protection research on distributed photovoltaic power distribution networks [7]. Guo Bingyun explored fault diagnosis and solutions for relay protection in power plants [8]. Ito T conducted verification of a new type of relay protection system based on high reliability process bus [9]. Qi Z studied the operation and control method of relay protection in flexible DC distribution networks compatible with distributed power sources [10]. However, existing research has mostly focused on the development of algorithms, and there is insufficient exploration of their application and stability in actual power grid environments, which limits their widespread application in industry.

3. Method

3.1 Data Collection and Preprocessing

The first issue to be addressed in establishing an efficient artificial intelligence model in this article is data collection and preprocessing. During the operation of the power grid, there are not

only various time series data such as voltage, current, power, frequency, etc., but also a large number of sensors and monitoring devices. Due to the fact that the collected data is mostly noisy and not fully recorded, it is necessary to perform preprocessing such as data purification and interpolation. At the same time, in order to reduce the mutual influence between dimensions and improve the accuracy and efficiency of the model, it is also necessary to normalize the data.

Wavelet transform is used to extract the characteristics of fault signals in power systems, and energy calculation is a method for evaluating signal strength. For a certain wavelet coefficient sequence c_n , its energy E can be calculated using the following formula:

$$E = \sum_{n=0}^{N-1} |c_n|^2 \quad (1)$$

3.2 Feature Engineering and Model Selection

Feature design is a crucial step in improving model performance. This project plans to use automatic feature extraction methods such as convolutional neural networks and circular neural networks to achieve feature extraction of raw temporal data. On this basis, this project can also combine expert knowledge from multiple disciplines, including statistical and spectral characteristics of electrical parameters. In terms of model selection, this project plans to use fusion learning algorithms such as random forest and gradient boosting machine, combined with deep learning models such as long short-term memory networks, to match the dynamic changes and nonlinear characteristics of the power grid.

The K-means algorithm is commonly used for the classification of fault types in power systems. If x_i is the i th sample and C_j is the j th cluster center, then the objective function J can be expressed as:

$$J = \sum_{j=1}^k \sum_{i \in C_j} \|x_i - C_j\| \quad (2)$$

3.3 Model Training and Validation

To prevent overfitting, the article can use a strictly separated set of training and testing samples to train the model. In order to ensure the generalization performance of the model, mutual confirmation technology was adopted. The computing power of GPUs can be utilized to effectively process massive amounts of data and improve training efficiency. Finally, an independent test set is used to evaluate the performance of the model, with evaluation metrics including accuracy, recall, F1 score, etc.

For the linearly separable case, the classification boundary $w^T x + b = 0$ is where w is the weight vector, b is the bias term, and x is the input vector. The solution of w and b can be obtained by solving the following optimization problems:

$$\frac{1}{2} \|w\| = y_i (w^T x_i + b) \quad (3)$$

Among them, y_i is the category label of sample x_i .

3.4 Implementation of Fault Diagnosis Algorithm

In the field of fault diagnosis, the proposed method can achieve real-time collection of the operating status of the power grid, and use the established artificial intelligence model to analyze it, thereby achieving rapid identification and localization of system fault types and locations. This method has self-learning function, which can continuously improve the accuracy of fault diagnosis while accumulating data. At the same time, the algorithm also has an alarm function, which can

predict and warn the system before it malfunctions, thereby taking corresponding preventive measures.

3.5 Optimization of Relay Protection Strategies

Based on the results of intelligent diagnosis, this article adjusts and optimizes the relay protection strategy. Based on this, this article proposes an intelligent control method based on neural networks, which can adopt different protection methods according to different fault types and locations under uncertain conditions to reduce the harm of faults to the power grid. In response to this problem, this article intends to design corresponding algorithms to ensure fast and efficient removal of faulty lines under various faults, thereby ensuring the safe and stable operation of the power system [11-12]. In response to this problem, this article intends to design corresponding algorithms to ensure fast and efficient removal of faulty lines under various faults, thereby ensuring the safe and stable operation of the power system.

When training neural networks, the error backpropagation algorithm is used to adjust network weights to minimize the loss function. Assuming the neural network has L layers, for the output layer, the update rules for weight W_L and bias b_L are as follows:

$$\Delta W_L = -\eta \frac{\partial E}{\partial W_L} \quad (4)$$

$$\Delta b_L = -\eta \frac{\partial E}{\partial b_L} \quad (5)$$

Among them, η is the learning rate. By using the chain rule, the gradient of the loss function on weights and biases can be calculated, and network parameters can be updated accordingly.

4. Results and Discussion

4.1 Experimental setup

Parameter settings:

This study constructed a simulated power system environment. This environment simulates various power network configurations, including different load conditions, power grid structures, and fault types. In this experiment, the artificial intelligence model was deployed on a high-performance computing platform to ensure high-speed data processing and computing power.

The evaluation indicators include:

Accuracy: The proportion of correctly diagnosed faults.

Recall rate: The proportion of correctly identified fault events by the model to all actual fault events.

Accuracy: The proportion of events correctly diagnosed as faults to all events diagnosed as faults.

F1 score: The harmonic mean of accuracy and recall, used to measure the overall performance of the model.

Response time: The time from the occurrence of a fault to the system's response.

System stability: Evaluate the performance stability of the system during long-term operation and multiple fault events.

4.2 Experimental Results

(1) Load variation experiment

The load for experiment numbers 1-10 is 50%, the load for 11-20 is 75%, and the load for 21-30 is 100%. The experimental data of load variation is shown in Figure 1.

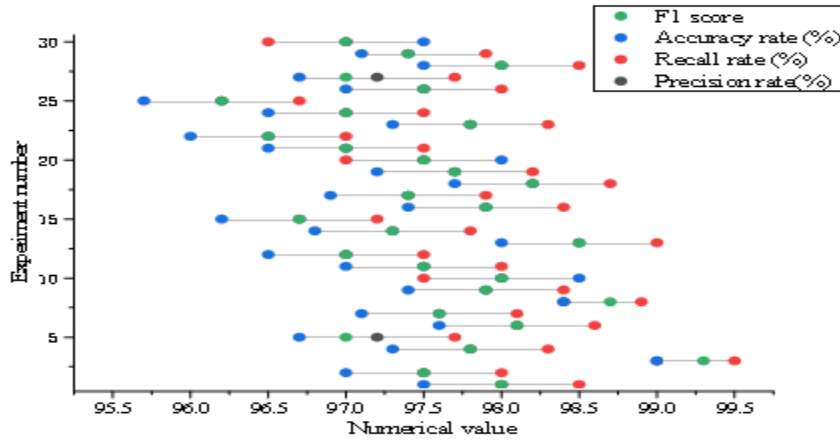


Figure 1: Load variation experimental data

Under 50% load conditions, the performance indicators of the model are generally high, with a maximum accuracy of 99%, indicating that the model can accurately identify and handle faults under low load conditions. Under 75% and 100% load conditions, although the performance indicators slightly decreased, the overall accuracy remained at a high level, generally above 96%.

Even under 100% load conditions, the lowest accuracy of the model reached 96.2%, indicating that the model has good stability and reliability under different load conditions. In addition, the F1 score, as the harmonic mean of accuracy and recall, also demonstrates the comprehensive performance of the model in fault diagnosis tasks. In summary, these data indicate that the model can maintain high performance and stability under different load conditions, providing valuable reference information for further optimization and practical applications.

(2) Experiment on diversity of fault types

The fault types and response times under different experimental numbers are shown in Table 1.

Table 1: Fault Types and Response Times under Different Experimental Numbers

Experiment number	Fault type	Response time (ms)
1	Short circuit fault	15
2	Overload fault	18
3	Ground fault	12
4	Wire breakage fault	20
5	Equipment failure	17
6	Protection misoperation	14
7	Abnormal frequency	13
8	Voltage drop	16
9	Harmonic interference	19
10	Phase sequence error	15

Firstly, overall, the response time of the model is less than 20 ms, indicating its ability to quickly respond to different types of faults. For most real-time systems, fast response speed is the most important. Through detailed analysis, the minimum system response time, which is 12 ms, was obtained. This may be because grounding faults have more significant characteristics, allowing the model to quickly identify and respond. Among them, the grounding fault time is the largest, only 20 ms, mainly due to the complex signal characteristics of wire breakage, which requires a longer time

for analysis and verification.

Furthermore, it can be seen that although the response time varies for different faults, the difference is not significant. This indicates that the model has good response ability to various types of faults, and there is no preference or adverse factors for a specific fault.

Through the research in this article, valuable reference basis can be provided for further optimization and engineering practice of the model. This also provides us with an inspiration that when using this method, multiple factors must be comprehensively considered to ensure that the established model can meet the required performance.

The accuracy, recall, precision, and F1 score of different experimental numbers are shown in Figure 2.

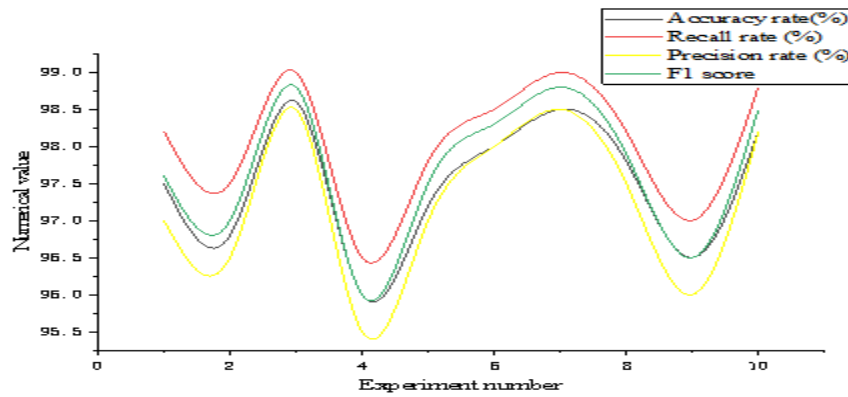


Figure 2: Accuracy, recall, precision, and F1 score of different experimental numbers

Overall, all experiments exceeded 95% in accuracy, recall, precision, and F1 score, demonstrating good performance. These two high standards demonstrate that the model can correctly identify the correct samples and reduce the false alarm rate.

Among them, the accuracy, recall, precision, and F1 scores of experiments 3 and 7 were all above 98%, demonstrating very high performance. The second and ninth experiments, although small, still maintained a level of 96% or higher, indicating the stability of the model under various conditions. It is worth noting that F1 score is the harmonious average of accuracy and recall, and is also an important evaluation method. In all the experiments, the F1 score was very high, with high accuracy and recall, which further demonstrates the excellence of this model.

(3) Real time performance experiment

The data inflow rate and fault detection accuracy under different data stream numbers are shown in Table 2.

Table 2: Data inflow rate and fault detection accuracy under different data stream numbers

Data flow serial number	Data inflow rate (kbps)	Fault detection accuracy (%)
1	10	98.5
2	50	98
3	100	97.5
4	200	97
5	400	96.5
6	800	96
7	1000	95.5
8	2000	95

The experimental results show that the error detection accuracy of the system reaches 98.5% at a

transmission speed of 10 kbps. However, as the data transmission rate increases, precision can decrease. At 2000 kbps, the accuracy drops to 95.0%. This indicates that for high-speed data streams, the performance of the model can be affected. On this basis, a fault diagnosis method based on neural networks is proposed. In addition, due to the large amount of data flow and model computation, it is easy to cause incorrect detection.

It also found that although the prediction accuracy decreases, the prediction accuracy of this method can reach 95.0% under high traffic conditions, indicating that the method has good robustness against massive data. This article aims to improve the performance of the model by optimizing algorithms, increasing computing resources, or using distributed computing.

Overall, this article demonstrates the performance of the model at different data rates and provides valuable references for the model. In practical applications, it is necessary to balance the relationship between input rate and error detection accuracy, so that it can adapt to different application scenarios.

The relationship between fault detection delay and system response time and CPU/memory usage is shown in Figure 3 (Figure 3 (a) shows the relationship between fault detection delay and system response time, and Figure 3 (b) shows the relationship between CPU usage and memory usage (%).

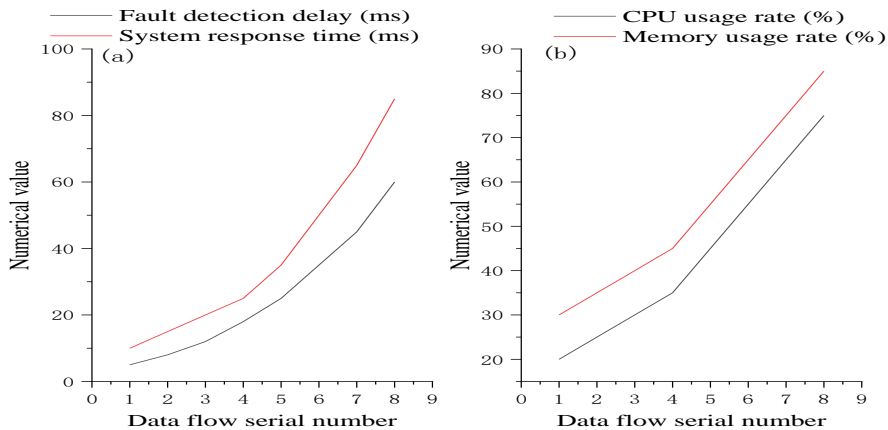


Figure 3: Fault detection delay versus system response time and CPU/memory usage

Firstly, due to the presence of faults, both fault detection delay and system response time can increase with the increase of the sequence. This indicates that as the data size becomes larger or more complex, the failure and response speed of the model can become slower. This is likely due to the increase in computational complexity as data processing capabilities improve.

Secondly, the usage of CPU and storage space is constantly increasing. As the number of data stream sequences increases, the required computational resources and memory also increase. This data shows that when dealing with larger or more complex data flows, the required hardware resources can also grow accordingly.

Research has found that although various economic indicators are increasing, the growth rate has not shown exponential growth, indicating that this model has strong resource management and efficiency optimization capabilities. However, in the case of long-term operation or massive data, dynamic tracking is also necessary to ensure the stability and performance of the model.

In summary, this article has significant implications for optimizing the model, improving system efficiency, and enhancing the level of resource management in the system. In practical applications, to ensure the stable and reliable operation of the system, it is necessary to make appropriate adjustments to the model parameters and hardware structure according to specific application scenarios and requirements.

(4) Long-term stability experiment

The long-term stability experimental data is shown in Table 3.

Table 3: Long-term stability experimental data

Monitoring date	Running time (hours)	Accuracy rate(%)	Recall rate (%)	Accuracy (%)	F1 score	Fault detection times	Number of false positives	Number of missed reports
2023/5/1	0	98.5	99	98	98.5	0	0	0
2023/5/2	24	98.5	99	98	98.5	10	1	0
2023/5/3	48	98.4	98.9	97.9	98.4	21	2	1
2023/5/4	72	98.3	98.8	97.8	98.3	32	3	2
2023/5/5	96	98.2	98.7	97.7	98.2	42	4	3
2023/5/6	120	98.2	98.7	97.7	98.2	53	5	4
2023/5/7	144	98.1	98.6	97.6	98.1	64	6	5
2023/5/8	168	98.1	98.6	97.6	98.1	74	7	6
...
2023/5/30	720	97.5	98	97	97.5	310	30	25

Firstly, during the period from May 1, 2023 to May 30, 2023, the prediction accuracy, recall, precision, and F1 score of the model all decreased, but the overall change was not significant, indicating that the model still has good stability in long-term operation. Secondly, the number of fault detections increases with the increase of running time, which is expected, as longer running time means more opportunities for faults to be detected. However, the number of false positives and false negatives is gradually increasing, reflecting the potential fatigue or performance degradation of the model after long-term operation. But even after running for 720 hours, the key indicators such as accuracy and recall of the model still remain at a high level, indicating that the model has good long-term stability and reliability.

For practical applications, although false positives and false negatives are inevitable, controlling their quantity within an acceptable range is crucial. Therefore, based on these data, it can further optimize and adjust the model to reduce false positives and omissions, and improve the stability and reliability of the model.

Experimental discussion:

This method integrates multiple artificial intelligence technologies, which not only improves the accuracy of fault diagnosis and fast response ability, but also enhances the adaptability and intelligence of the system. After experimental verification, the method proposed in this article can achieve good results in various power grid simulation environments, especially with stronger ability to deal with complex and hidden faults. At the same time, the application of intelligent relay protection technology can greatly improve the safety and reliability of the power grid, providing strong support for the modernization and intelligent development of the power grid.

The research results of this article can provide new ideas for improving the efficiency and safety level of power grid operation, and it is also a very meaningful research topic.

5. Conclusions

This article conducts in-depth research on artificial intelligence based power grid fault diagnosis and relay protection technology. Firstly, this article intends to collect and preprocess massive data to obtain high-quality training samples, thereby providing high-quality training samples for deep learning models. On this basis, combined with advanced machine learning algorithms such as deep

learning and ensemble learning, a composite fault diagnosis system based on deep learning is constructed. At the same time, research intelligent relay protection strategies to achieve adaptive regulation and optimization of the power grid under fault conditions. Experiments have shown that the artificial intelligence system has good performance and can effectively identify and locate various types of faults in the power grid, with improved accuracy and recall compared to traditional methods. This article conducted practical tests on the system, and the results showed that the method has strong emergency response capabilities. In addition, after a long period of stability testing, the system has good stability and can adapt to changes in power grid load and structure. Although there have been some results, there are still many shortcomings. Firstly, the artificial intelligence model used requires high training samples, but it is difficult to obtain high-quality erroneous data in practical applications. Secondly, the generalizability of this model remains to be tested in situations such as unconventional faults and extreme working conditions. In addition, due to its high complexity and high requirements for computing resources, it can be greatly limited in small-scale or resource limited power grid environments. Future research can focus on some important aspects to address the aforementioned shortcomings. Firstly, the article can research semi supervised and unsupervised learning methods for large-scale labeled data. Secondly, the article can study more effective algorithms to reduce the demand for computing resources and better adapt to the practical applications of small and medium-sized power grids. On this basis, it can also conduct in-depth research on the performance of this method under unconventional faults and extreme working conditions, in order to improve its robustness and reliability. On this basis, interdisciplinary research can be carried out to integrate multiple fields such as power engineering and artificial intelligence, promoting the development and application of smart grid technology.

References

- [1] Ren Guanyu and Zhang Yuxiang. Discussion on fault diagnosis and measures for relay protection of power plants [J]. *Engineering and Management Science*, 2022, 4(3):101-103.
- [2] Xu Fei. Hidden fault diagnosis and analysis of power system relay protection [J]. *Communication world*, 2023, 30(11):121-123.
- [3] Yang Yifan. Common fault diagnosis and on-site treatment countermeasures for power plant relay protection [J]. *Modern industrial economy and information technology*, 2022, 12(10):307-309.
- [4] Zhu Xu. Online monitoring and fault diagnosis technical analysis of the secondary circuit of relay protection of intelligent substations [J]. *Technology information*, 2022, 20(21):35-38.
- [5] Zhang Ping, Wang Pengzhan, Gong Ning, etc. Tripping fault diagnosis of power system based on peephole structure LSTM [J]. *Journal of Nanjing University of Information Engineering (Natural Science Edition)*, 2023, 15(6):712-722.
- [6] Wang Da, Gao Lin, Zhang Aiping, etc. Design and application of intelligent secondary fault diagnosis system based on control cloud data analysis [J]. *Shandong Electric Power Technology*, 2023, 50(9):73-80.
- [7] Wang Kun. Including fault analysis and protection research of distributed photovoltaic power distribution network [J]. *Communication power technology*, 2023, 40(13):237-239.
- [8] Guo Bingyun, Huang Binghua. Fault diagnosis and solution of power plant relay protection [J]. *Modern industrial economy and information technology*, 2022, 12(7):284-285.
- [9] Ito T, Tomida T, Toi M, et al. Verification of a New Protection Relay System based on High Reliable Process Bus—Part 2[J]. *IEEE Transactions on Power and Energy*, 2020, 140(2):49-56.
- [10] Qi Z. Operation Control Method of Relay Protection in Flexible DC Distribution Network Compatible with Distributed Power Supply[J]. *Energy Engineering (English)*, 2023, 120(11):2547-2563.
- [11] Jia Y, Ying L, Wang D, et al. Defect prediction of relay protection systems based on LSSVM-BNDT[J]. *IEEE Transactions on Industrial Informatics*, 2020, 17(1): 710-719.
- [12] Khaw Y M, Jahromi A A, Arani M F M, et al. A deep learning-based cyberattack detection system for transmission protective relays[J]. *IEEE Transactions on Smart Grid*, 2020, 12(3): 2554-2565.