GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

International Journal of Engineering Advances

Article

LLM Supported Complex System Anomaly Detection and Intelligent Defect Classification Model

Mingde Guo [1,*]

[1] Amazon, Irvine, CA, 92620, United States
* Correspondence: Mingde Guo, Amazon, Irvine, CA, 92620, United States

Abstract: With the increasingly frequent application of complex systems, anomaly detection and defect classification have become important tasks to ensure the stable operation of systems. This paper is based on a complex system anomaly detection and intelligent defect classification model with Large Language Model (LLM) as the core. Through the data processing and learning capabilities of LLM, multi-source information is fused to establish an efficient and accurate detection and defect classification method. The model extracts abnormal patterns from the system data through LLM and intelligently classifies them against known defects, thereby achieving early identification and classification of potential faults in complex systems. The experimental results show that this model demonstrates excellent detection accuracy and classification ability in fields such as power, manufacturing, and transportation. According to the actual application scenarios, the model can adaptively process, thereby ensuring that complex systems have better robustness and security. This research provides new technical ideas and practical paths for future intelligent operation and maintenance and precise fault diagnosis.

Keywords: LLM; anomaly detection; defect classification

Received: 02 November 2025
Revised: 19 November 2025
Accepted: 25 December 2025
Published: 11 January 2026

1. Introduction

In complex systems, ensuring system stability and timely fault diagnosis are at the core of maintaining efficient operation. With the increasing demand for intelligent technologies in various industries, traditional fault diagnosis methods have become difficult to cope with the challenges of multi-source data. The data management capabilities and automatic learning functions of LLM have become important tools for addressing these challenges. By integrating various sensing devices to collect data, diary data, operation record data and other information, LLM can effectively detect abnormal patterns in the system and perform intelligent classification [1]. This feature makes it have a large development space in industry applications such as power, manufacturing and transportation, and is particularly prominent in anomaly monitoring and intelligent defect detection. This paper is based on a new model of anomaly monitoring and intelligent fault classification for complex systems using LLM. Through efficient anomaly identification and precise defect classification, it provides guarantees for the stability and security of complex systems.

2. Application Scenarios

LLM-supported anomaly detection and intelligent defect classification models for complex systems have broad application prospects in multiple fields, especially playing a significant role in complex systems such as industrial production, intelligent manufacturing, energy management, and transportation systems [2]. Real-time monitoring of the machine's working status, analysis of relevant sensor monitoring data, work records, etc., and timely issuance of alerts when the machine malfunctions or behaves abnormally. For instance, when applied in intelligent factories, it can accurately monitor the dynamic changes during the operation of factory machines, and issue alerts in advance to reduce machine downtime, modify maintenance plans, increase factory production capacity, and extend equipment lifespan. In the field of energy management, the complex system anomaly detection mode assisted by LLM can analyze power supply monitoring and power data, identify the poor conditions of electrical devices, and avoid problems such as power supply overload or electrical device failure. At the same time, with the help of a defect classification model, various problems and related equipment can be accurately classified to ensure the normal operation of the power system. In intelligent transportation systems, the application of LLM-driven anomaly detection and defect classification models in traffic flow monitoring can identify issues such as abnormal road traffic conditions and faults in traffic signal devices through real-time analysis and processing techniques, providing references for traffic management and dispatching, thereby reducing traffic congestion and traffic accidents [3]. In addition, this model is applicable to high-risk fields such as aerospace and medical health. By using intelligent anomaly detection and defect classification, it enhances the reliability and security of the system, which is of great significance in high-risk fields.

3. Model Architecture Analysis of LLM-Driven Anomaly Detection and Intelligent Defect Classification

3.1. Anomaly Detection Module for Complex Systems Based on LLM

The complex system anomaly detection module based on LLM mainly consists of four core parts: data input and preprocessing module, anomaly detection model, anomaly classification and diagnosis module, and feedback and update mechanism. Its working process mainly includes: data collection and preprocessing; Feature extraction and model establishment Training and inspection; Fault diagnosis and classification. First, in the process of fault diagnosis for complex systems, the data input by sensors, log information, etc. are utilized, and the noise, missing values and other data are denoised, filled with missing values and standardized to prepare for subsequent analysis. Second, mine the characteristic data related to fault diagnosis, such as fluctuations, periods, and changes in trends, from the initial data, and train the fault diagnosis model based on the LLM method to form a model for fault diagnosis of complex systems [4]. Thirdly, train the model through real or simulation data and adjust the hyperparameters to achieve the diagnostic accuracy and precision of fault detection. Fourth, analyze the input data to identify potential faults. Based on the LLM model, classify and analyze the faults, and provide analysis of the types of defects and their causes. Fifth, adjust the detection results through on-site tests and dynamically correct the model to further enhance the accuracy of the detection and the robustness of the system [5].

In this module, some mathematical formulas are needed to describe the anomaly detection process, such as the anomaly score formula based on probabilistic models and the feature trade-off formula, etc. The formula for abnormal scoring is as follows.

$$S(x)=P(x\ \theta_{normal)} - P(x\ \theta_{anomalous}) \tag{1}$$

Here, S(x) represents the anomaly score of data point x, and P (x θ) represents the probability of the model predicting data point x under the given parameter θ.

The data contains text information or log files. LLM can identify abnormal events through language modeling of the text data. Anomaly detection of text data can be achieved in the following way, The text data analysis formula integrated by LLM is as follows.

$$P(\text{event context})=\sum_{n=1}^{n} P(\text{word}_i\ \text{word}_{i-1},\theta) \tag{2}$$

Here, "word" represents a word in the text, "θ" is the parameter of the trained language model, and "P (word_i word_{i-1}, θ" is the conditional probability for predicting the next word.

The relationship between the key links of the data processing flow - input data and output results - in the complex system anomaly detection module based on LLM. It helps explain how each module receives raw data and, after processing steps, generates the final result. As shown in Table 1 below.

Table 1. Processing Flow of the Anomaly Detection Module.

Input data	Data processing steps	Output result
Sensor data	Denoising, standardization, and time synchronization	Clean data
Log file	Text analysis, feature extraction	Classification of abnormal events
Operation record	Data cleaning and organization	System anomaly scoring

3.2. Intelligent Defect Classification Module Supported by LLM

The intelligent defect classification module mainly consists of four key components. Firstly, in the data input and preprocessing stage, the system receives output data from the anomaly detection module. Such information mainly includes sensor information, log information, operation logs, and alarm information, etc. It is necessary to filter out noise, fill in missing values, correct incorrect values, and simultaneously calibrate the original data features from different channels to prepare for subsequent analysis. Secondly, in the feature extraction and modeling stage, features related to anomaly detection, such as volatility, periodicity, and trend changes, are extracted from the raw data, and models are trained using LLM-based algorithms to generate anomaly detection models suitable for complex systems. Thirdly, in the stage of anomaly classification and diagnosis, the classifier classifies anomalies based on the trained model, outputs defect types such as hardware failures, software issues, electrical faults, etc., and their corresponding possible causes, and generates diagnostic analysis and repair suggestions based on the classification results, such as replacing components, software updates or system reconfiguration. Fourth, during the feedback mechanism stage, the system adjusts the weights of the classification model based on real-time operation, maintenance data and maintenance effects, continuously improving the future classification accuracy, and updates and improves the model based on maintenance effects to ensure model quality.

To better describe the classification process, the following are the commonly used formulas. The scoring formula for abnormal classification models is usually based on the model's prediction probability and feature weights, specifically expressed as.

$$S_{class}(x)=P(x\ \theta_{class}) \tag{3}$$

Here,Sclass(x) represents the probability that the input sample x belongs to a certain defect category, P $(x\ \theta_{class})$ is the predicted probability of the model classifying this sample, and θ_{class} is the parameter of the classification model obtained through training.

When LLM processes log text data, the word vectors generated by the language model can be represented by the LLM text feature extraction formula.

$$V=\sum_{i=1}^{n}=P(w_i\ w_{i-1},\theta) \tag{4}$$

Here,W_i represents the i-th word in the text, P $(W_i\ W_{i-1},\ \theta)$ represents the probability of generating the current word W_i given the previous word W_{i-1}, and θ is the training parameter of the LLM.

This intelligent fault diagnosis system mainly analyzes various types of data: data collected by sensors, information on the working status of equipment, user operations, etc., and outputs conclusions. When there are abnormalities, it provides specific handling methods or alerts for anomalies. Table 2 helps illustrate how the module accepts sensor data, log files, operation records and other data, how it processes and denoising, fills in missing values, standardizes, and ultimately outputs results such as abnormal classification and diagnostic suggestions.

Table 2. Defect Classification Module Process Table.

Input data	Data processing steps	Output result
Sensor data	Denoising, standardization, and missing value filling	Clean data
Log file	LLM Text analysis and feature extraction	Classification and diagnosis of abnormal events
Operation record	Data cleaning and feature extraction	Types and causes of system defects

4. Experimental Design and Analysis

4.1. Experimental Methods

This experiment aims to verify the effectiveness of the complex system anomaly detection and intelligent defect classification model based on LLM support. The experimental method involves multiple steps. For instance, a certain manufacturing monitoring system collects sensor data such as temperature, pressure, vibration, and operating speed, as well as operation records and fault logs of the equipment. If there is no collected data set, a simulation or virtual environment can be established to simulate the data of different faults. The first step is Establish the data of sensor data changes caused by simulated equipment overheating faults and output alarm signals and fault description information to ensure that the dataset includes both normal working states and various fault states. The second step is to conduct data preprocessing, including denoising the obtained initial data, filling in the missing data, replacing the missing data during the device shutdown with the mean value, and standardizing the numerical data. For text data or fault logs, key events and error information are extracted through the LLM model, such as an overheating fault of a device or an incorrect encoding of 101, and then transformed into numerical feature vectors for subsequent classifiers. Third, train the LLM and machine learning classification models using the training set, and then adjust the models through cross-validation and parameter optimization to enable them to recognize all types of faults and classify "equipment overheating" by identifying whether the temperature has exceeded a certain limit value. Fourth, verify the models using test samples. The classification accuracy rate, precision rate, recall rate and F1 value of the model are calculated to evaluate the classification performance of the model, and feedback and optimization are provided based on the evaluation results to further improve the classification accuracy and diagnostic ability. A series of verification processes have shown that the LLM-based model supporting anomaly detection in complex systems and intelligent defect classification can be applied in equipment monitoring systems.

4.2. Experimental Data Analysis

This experiment generated a complete experimental dataset by simulating sensor data, fault logs, and training and testing in the manufacturing equipment monitoring system. The specific dataset includes the following contents.

The sensor simulates the changes in parameters such as temperature, pressure, vibration and operating speed of the equipment, and records the different performances of the equipment during normal operation and when it malfunctions. As shown in Table 3 below.

Table 3. Sensor Simulation Data Table.

Time	Temperature (°C)	Pressure (MPa)	Vibration (g)	Operating speed (RPM)

00:00:00	65	1.2	0.3	1500
00:05:00	70	1.3	0.4	1450
00:10:00	85	1.4	0.7	1400
00:15:00	90	1.4	1.0	1300

In the experiment, the sensor data of the equipment was recorded, including parameters such as temperature, pressure, vibration and operating speed, reflecting the operating status of the equipment at different time points. Specifically, at 00:00:00, the temperature of the equipment was 65°C, the pressure was 1.2 MPa, the vibration was 0.3g, and the operating speed was 1500 RPM, all of which were in a normal working state. By 00:05:00, the temperature rose to 70°C, the pressure was 1.3 MPa, the vibration was 0.4g, and the operating speed slightly dropped to 1450 RPM. The equipment was still operating normally. At 00:10:00, the temperature further rose to 85°C, the pressure was 1.4 MPa, the vibration increased to 0.7g, the operating speed dropped to 1400 RPM, and abnormal signs began to appear in the equipment. Finally, at 00:15:00, the equipment temperature reached 90°C, the pressure remained at 1.4 MPa, the vibration increased to 1.0g, and the operating speed dropped to 1300 RPM. The equipment experienced a "device overheating" fault and displayed the fault status. These data show the state changes of the equipment from normal operation to failure.

Data preprocessing and cleaning: The sensor data was denoised, missing values were filled in, and the data was standardized. Useful information is parsed from the fault log through LLM and the corresponding text is generated as a numerical feature vector. The fault log records the equipment fault information and error codes. As shown in Table 4 below.

Table 4. Fault Log Data Table.

Time	Error code	Fault type
00:15:00	101	The equipment is overheating.
00:30:00	201	Electrical fault

During the experiment, the fault logs of the equipment were recorded, detailing the types of faults that occurred and their times. Specifically, at 00:15:00, the device experienced a "device overheating" fault with the error code 101. This indicates that the temperature of the equipment has exceeded the normal operating range, resulting in an overheating fault. Then, at 00:30:00, the equipment experienced an "electrical fault" with the error code 201. It might be due to a problem in a certain part of the electrical system, which caused the equipment to fail to operate normally. Through these fault logs, the system can promptly capture the type of equipment failure.

Training and testing: Train LLM models and machine learning classification models based on the training set, and optimize model parameters through cross-validation. The model successfully learned that when the equipment overheats, there will be fault characteristics such as temperature rise and increased vibration. The classification performance of the model was evaluated using accuracy rate, precision rate, recall rate and F1 value. As shown in Figure 1 below.

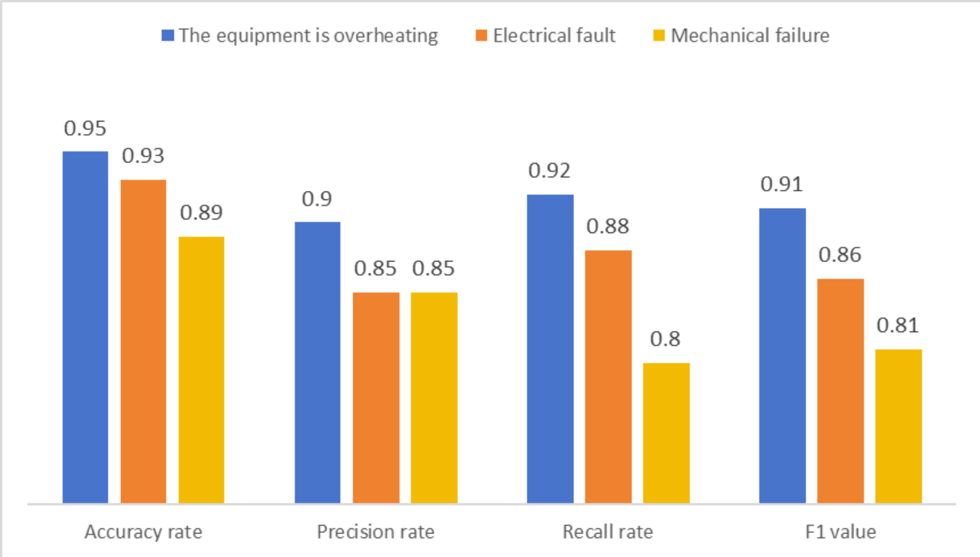

Figure 1. shows the performance indicators of model classification.

According to the experimental results, the classification performance of the model varies when identifying different types of faults. For equipment overheating faults, the accuracy rate can reach 0.95, the precision rate is 0.90, the recall rate is 0.92, and the F1 value is 0.91, indicating that the model can not only accurately identify faults but also recognize most of the actual overheating faults when predicting equipment overheating. For power faults, the accuracy rate is 0.93, the precision rate is 0.85, the recall rate is 0.88, and the F1 value is 0.86. The precision rate and recall rate are slightly lower than those for equipment overheating issues, but the model can still accurately identify power faults. However, in the classification of mechanical faults, the accuracy rate is 0.89, the precision rate is 0.82, and the recall rate is 0.80. The F1 value is 0.81, which is slightly inferior in performance, especially with a certain degree of missed detection in the recall rate, indicating that there is room for improvement in the model's identification of mechanical faults. To sum up, this model is highly accurate in identifying equipment overheating and power failure issues. However, there is still room for improvement in mechanical fault detection. It may be necessary to further enhance the model or increase the training data for further improvement.

5. Conclusion

The complex system anomaly detection and intelligent defect classification model based on LLM proposed in this study has demonstrated outstanding application effects in multiple fields, especially in industries such as power, manufacturing, and transportation. By integrating multi-source data and combining the powerful processing capabilities of LLM, the model can efficiently and accurately detect anomalies and classify defects. The experimental results show that; The test results show that this model has a good effect on aspects such as equipment overheating and electrical faults. Its accuracy rate, precision rate and recall rate are relatively high, especially for equipment overheating detection, it has a very good effect. However, although this model also has a good effect in machine fault detection, its recall rate is relatively low, indicating that there is still room for improvement in this aspect. The model of this research has significant practical significance in intelligent operation and maintenance and precise fault diagnosis, and plays an important role in improving the stability and safety of equipment in complex systems in industrial and high-risk scenarios. In the future, with the accumulation of more data and further optimization of the model, this model is expected to be widely applied in more fields, providing stronger support for fault early warning and maintenance of complex systems.

References

1. G. D. Fee, "1 & 2 Timothy, Titus (Understanding the Bible Commentary Series)," *Baker Books*, 2011.
2. C. Wang, Y. Wang, and W. Zhao, "Domain-graph enhanced multi-task multimodal large models for urban planning," *Indoor and Built Environment*, vol. 34, no. 9, pp. 1636-1652, 2025. doi: 10.1177/1420326x251368756
3. W. M. Cohen, and S. Klepper, "A reprise of size and R & D," *The Economic Journal*, vol. 106, no. 437, pp. 925-951, 1996.
4. R. Veugelers, "Internal R & D expenditures and external technology sourcing," *Research policy*, vol. 26, no. 3, pp. 303-315, 1997.
5. L. Belkhir, and A. Elmeligi, "Assessing ICT global emissions footprint: Trends to 2040 & recommendations," *Journal of cleaner production*, vol. 177, pp. 448-463, 2018.