

LOG x: AI-Powered Log Analysis and Automated Incident Response Using Large Language Models

Yash Paraskar¹, Sushrut Morde², Aarohi Gholve³, Ziya Attar⁴, Nitish Das⁵

¹²³⁴ Student Authors, Department of Computer Science MIT ADT University, Pune, India

⁵ Faculty Advisor, Department of Computer Science MIT ADT University, Pune, India

Abstract - Log analysis is an important part of keeping an eye on systems, keeping them safe, and working in DevOps. But traditional methods depend a lot on manual inspection and rule-based systems, which are not very efficient and can't be changed. This paper introduces LOG x, an AI-driven log analysis platform that employs extensive language models to automate the classification of logs, the identification of root causes, and the creation of incident reports. The system uses a quantized Gemma 2B model to make inference faster and lets you see and report on structured PDFs. Experimental observations indicate that the proposed system enhances both accuracy and efficiency relative to traditional methods, rendering it suitable for implementation in contemporary computing environments.

Key Words: Log analysis, Large Language Models, Incident Response, DevOps, Root Cause Analysis, Model Quantization

1. INTRODUCTION

Modern software systems constantly make huge amounts of logs that are very important for keeping an eye on how the system works, finding bugs, and spotting possible security threats. As systems become more complicated, the amount and variety of log data grow a lot, making it hard to analyze by hand.

Engineers and system administrators often spend a lot of time trying to figure out what logs mean, which slows down incident response and makes it more likely that important problems will be missed. Traditional log analysis methods depend on rules that have already been set and pattern matching techniques. These methods don't work well with systems that change and grow over time. These methods don't work well when log messages have relationships with other messages, and they don't work well when they encounter new situations.

This paper presents LOG x, an intelligent log analysis system that utilizes large language models to automate log classification, root cause analysis, and incident report generation. The system is made to be easy to use, make things more efficient, and give you useful information that you can act on.

2. RELATED WORK

People have done a lot of research on log analysis in the areas of system monitoring and cybersecurity. Early methods used rule-based systems that relied on patterns that people had to define by hand to find mistakes and problems. These methods work well for problems that are already known, but they aren't very flexible and need to be updated all the time to stay useful. Later, machine learning methods were added to make systems more adaptable by finding patterns in old data. These models were somewhat successful at classifying and finding anomalies, but they often needed a lot of feature engineering and labeled datasets. Deep learning methods improved performance even more by automatically learning representations from raw data. However, they still had trouble with understanding and reasoning in context. The rise of large language models has greatly changed how we do natural language processing by making it possible to understand and reason in context. These models offer a promising solution for log analysis by interpreting logs in a way that is similar to how people think, which gets around the problems with traditional methods.

3. SYSTEM ARCHITECTURE

LOG x's architecture is a layered pipeline that makes sure log data is processed quickly and can be scaled up. The system starts with an input interface where users can send in log data as plain text or in structured formats like CSV files.

A preprocessing module takes this input and cleans and normalizes it to get rid of inconsistencies and get it ready for analysis. Then, the processed data is sent to the core inference engine, which is based on a large language model. The model looks at the logs and figures out what they mean in context. It also finds patterns that show what kind of events are happening and how bad they are. This method doesn't rely on preset rules like traditional systems do, so it can change to fit new and unknown log patterns.

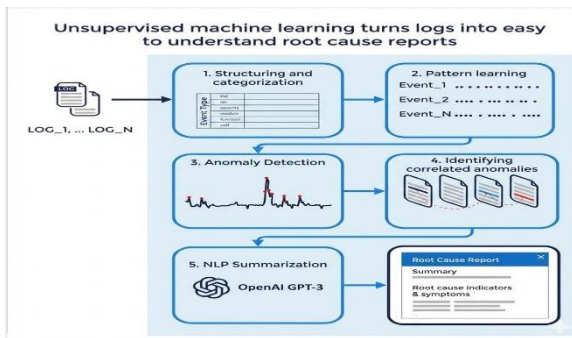


Fig. 1. LOG x System Architecture

An interpretation module then processes the analysis results and pulls out important information, such as classification, root causes, and suggested fixes. Finally, the output is shown on a web page, and the system lets you make structured reports in PDF format.

4. System Workflow

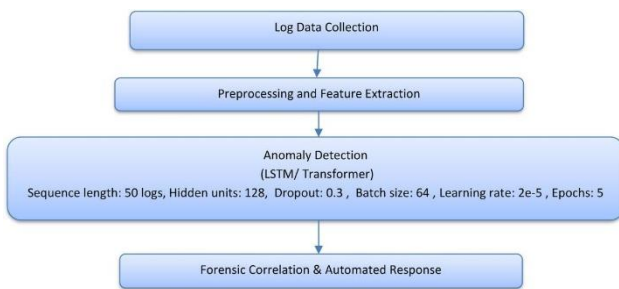


Fig. 2. Workflow of LOG x

The workflow of LOG x is a step-by-step process that turns raw log data into useful information. At first, the user enters information through the interface, and then the system processes the data to make sure it is consistent. The large language model then analyzes the cleaned logs, doing semantic interpretation and classification. By looking at how the data is related to each other, the model figures out what type of each log entry is and what might have caused it. After that, the results are put together in a way that makes sense, with explanations and possible fixes. The system combines the information from multiple logs to make a complete incident report. The workflow also has visualization tools that let users see trends and patterns in log data.

5. METHODOLOGY

The methodology used in LOG x is based on using large language models' ability to understand context. The first step is preprocessing, which cleans up the logs and makes them consistent by getting rid of noise and other problems. This step makes sure that the data that was entered can be analyzed. After the logs have been cleaned, they are sent to

the language model, which uses inference to look at the text in each entry. The model finds patterns, relationships, and oddities that show what kind of log it is. The system uses this analysis to give the problem a classification label and a natural language explanation. Root cause analysis looks at dependencies and contextual clues in the logs to figure out what might have gone wrong. The model also makes troubleshooting steps that give you useful advice on how to fix problems. This method does away with the need to write rules by hand and lets the system adapt to new log patterns on the fly.

6. IMPLEMENTATION

Python and the Flask framework are used to build the backend of the LOG x system. Using transformer-based architectures, large language models can be combined. PyTorch is the framework that does the math behind it all. We use efficient libraries that can handle large amounts of log data to do data processing tasks. Model quantization techniques are used to make things work better and use less computing power. Using 4-bit quantization cuts down on memory use a lot while still keeping accuracy at an acceptable level. The system also has modules for making reports and visualizations, which make it easier to use and give users a full picture.

7. EXPERIMENTAL RESULTS

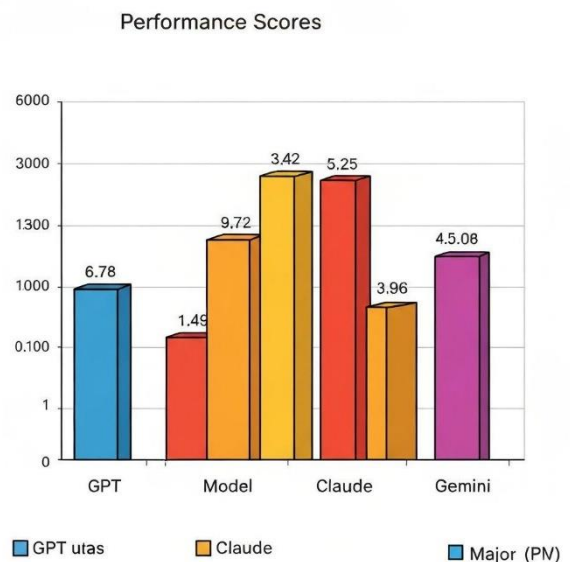


Fig. 3. Performance Comparison

We used a dataset with different types of log entries, such as errors, warnings, and informational messages, to test how well LOG x worked. The assessment concentrated on classification precision and the caliber of derived insights. The results show that the suggested system is very good at telling the difference between different types of logs. Using large language models helps us better understand the

meaning of logs, which leads to more accurate classification than traditional methods. Also, the explanations and troubleshooting steps that were made were found to be useful and relevant in real-life situations. The system also strikes a balance between efficiency and performance. While large language models consume higher computing resources, the use of quantization methods make

[5] A. Vaswani et al., "Attention Is All You Need," *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[6] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, "QLoRA: Efficient Finetuning of Quantized LLMs," *arXiv preprint arXiv:2305.14314*, 2023.

8. LOG VISUALIZATION

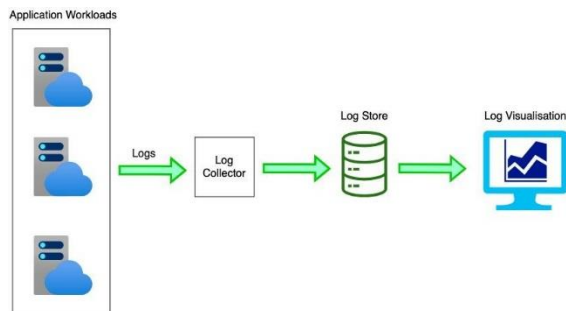


Fig. 4. Log Visualization Dashboard

LOG x visualization module helps analyze log trends in a very intuitive manner. Using the data provided, LOG x will create visual representations, which help identify how frequent certain logs occur. This makes it easy for a user to spot errors and any kind of abnormal behavior in the system.

9. CONCLUSIONS

In this research paper, LOG x was developed as an AI based tool that can be used for automating log analysis and incident responses. The application of large language models enables the system to eliminate the shortcomings of previous systems based on rules and offers more intelligent solutions by providing meaningful explanations and insights. The findings show that LOG x can effectively contribute to the improvement of both accuracy and user experience. Further development of the system is expected to address issues related to its real-time nature and enterprise integration.

REFERENCES

[1] Hugging Face, "Transformers Documentation."
 [2] Py Torch, "Py Torch Documentation and Tutorials."
 [3] Google, "Gemma: Open Models Based on Gemini Research and Technology."
 [4] S. He, J. Zhu, P. He, and M. R. Lyu, "Experience Report: System Log Analysis for Anomaly Detection," *IEEE 27th International Symposium on Software Reliability Engineering (ISSRE)*, 2016.