Compliance Monitoring and Enforcement through Log Analysis using Large Language Models (Infosec Engineering)

Problem Statement :

As Flipkart deals with increasing volumes of data and complex systems, ensuring compliance with security policies, standards and baselines has become a critical challenge. To address this issue, we propose a project focused on developing a system that leverages large language models for compliance monitoring and enforcement through log analysis from relevant sources.

The objective is to build a solution that can effectively analyze logs, system configurations, access controls, and user privileges to check for compliance with security policies and standards. By utilizing the power of large language models (LLMs) like ChatGPT or its open source alternatives, we aim to automate the process of identifying non-compliant activities and generating actionable insights for remediation.

—----------

Here are the following attributes that we have to keep in regard while developing this.

Solution Attributes of Compliance Monitoring and Enforcement :

- **Rule Definition**: The ability to distill the rules and use the standards provided to infer the possible relationships between rules and parameters in the logs/policies.
- **Flexibility**: The ability of the system to handle various log formats (CSV, text, PDF), rule sets, and compliance standards. It should be adaptable to different environments and configurations.
- **Actionable Insights**: The relevance and usefulness of the actionable insights provided by the system for *remediation* purposes
- **System Performance:** The efficiency and scalability of the system in handling large volumes of logs and text data.
- **Adaptability**: The system's capability to learn from new rules, update compliance standards, and incorporate feedback for continuous improvement without restarting
. ● **High Accuracy/Precision/Recall**: The ability to identify compliance breaches along with low false positives and and very low false negatives.
- **UI/UX (Optional):** a way to interact with the system, configure rules, upload logs and text files, view compliance reports, and access actionable insights.

Application workflow is that we provide some compliance documents as ruleset along with logs/system policies etc with known failures. Then we should be provided with the compliance breaches along with citations to the specific line/log file/userid. The system should also provide actionable insights i.e. what can be done to fix the issues.

Technical Requirements - An open source LLM (APIs not permitted because they send data to third party servers) + generic tech stack for automation and parsing.