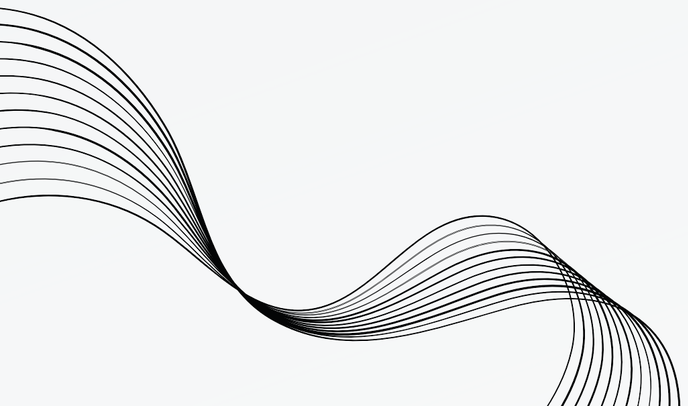
STEALTH BOT

PROBLEM STATEMENT

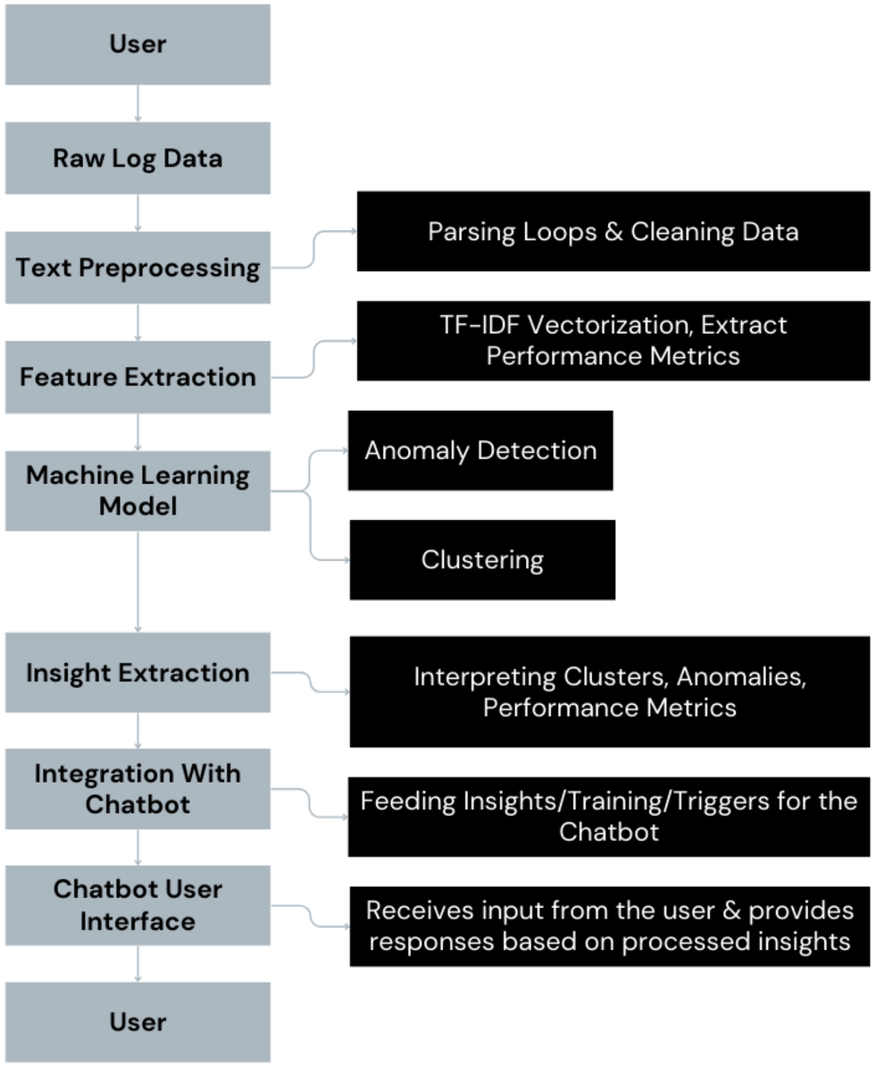
Generate insights from Linux log files using AI&ML tools and Generative AI techniques to query requests in natural language using LLM models



PROPOSED SOLUTION

We use Linux log file insights to improve our Language Model (LLM). It converts these insights into graphs and pie charts, summarizing findings. Users can directly upload their log files to our Chatbot for analysis.

WORKFLOW


User
Raw Log Data
Text Preprocessing
Parsing Loops & Cleaning Data
Feature Extraction
TF-IDF Vectorization, Extract Performance Metrics
Machine Learning Model
Anomaly Detection
Clustering
Insight Extraction
Interpreting Clusters, Anomalies, Performance Metrics
Integration With Chatbot
Feeding Insights/Training/Triggers for the Chatbot
Chatbot User Interface
Receives input from the user & provides responses based on processed insights
User

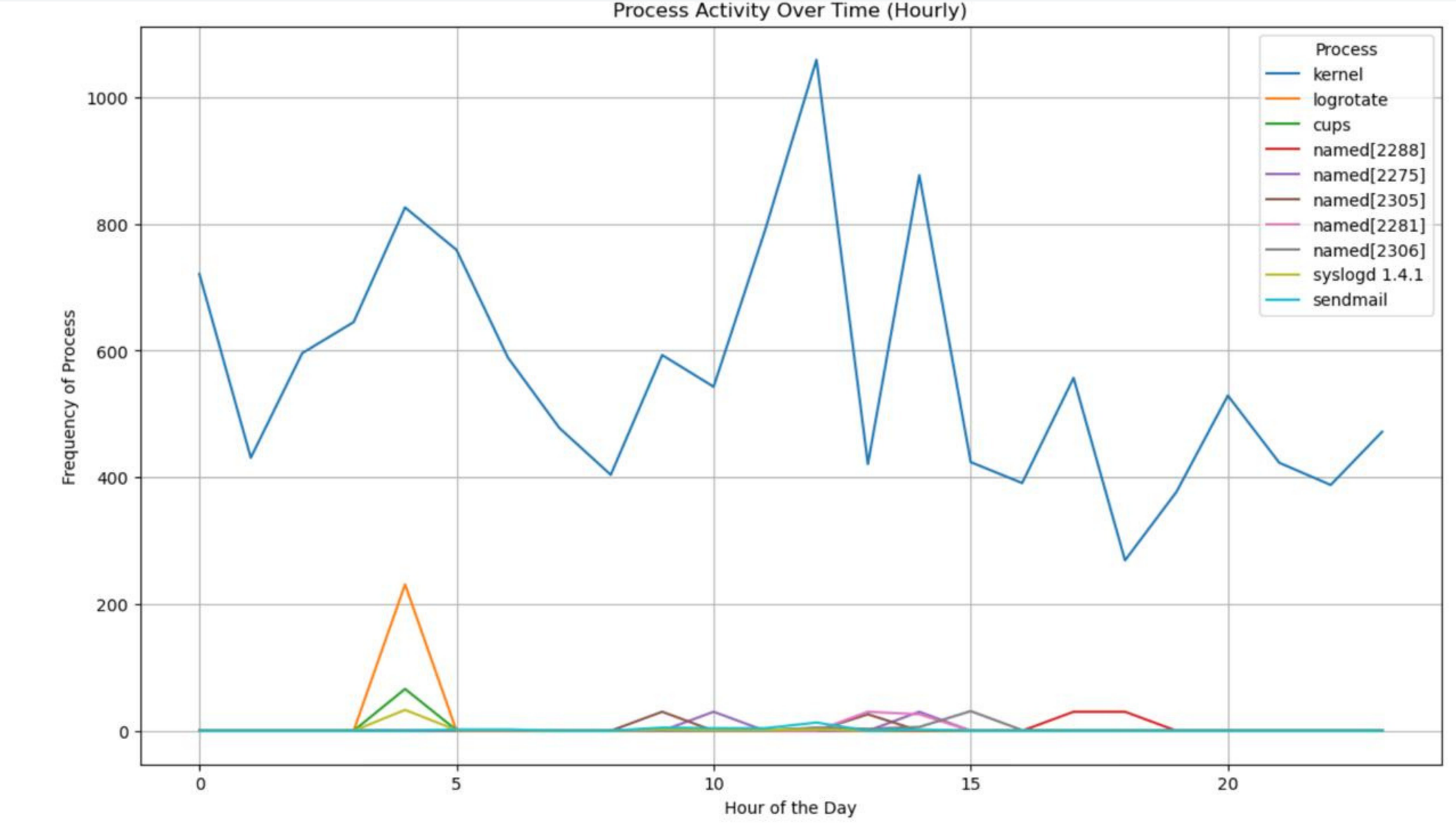
UNIQUENESS

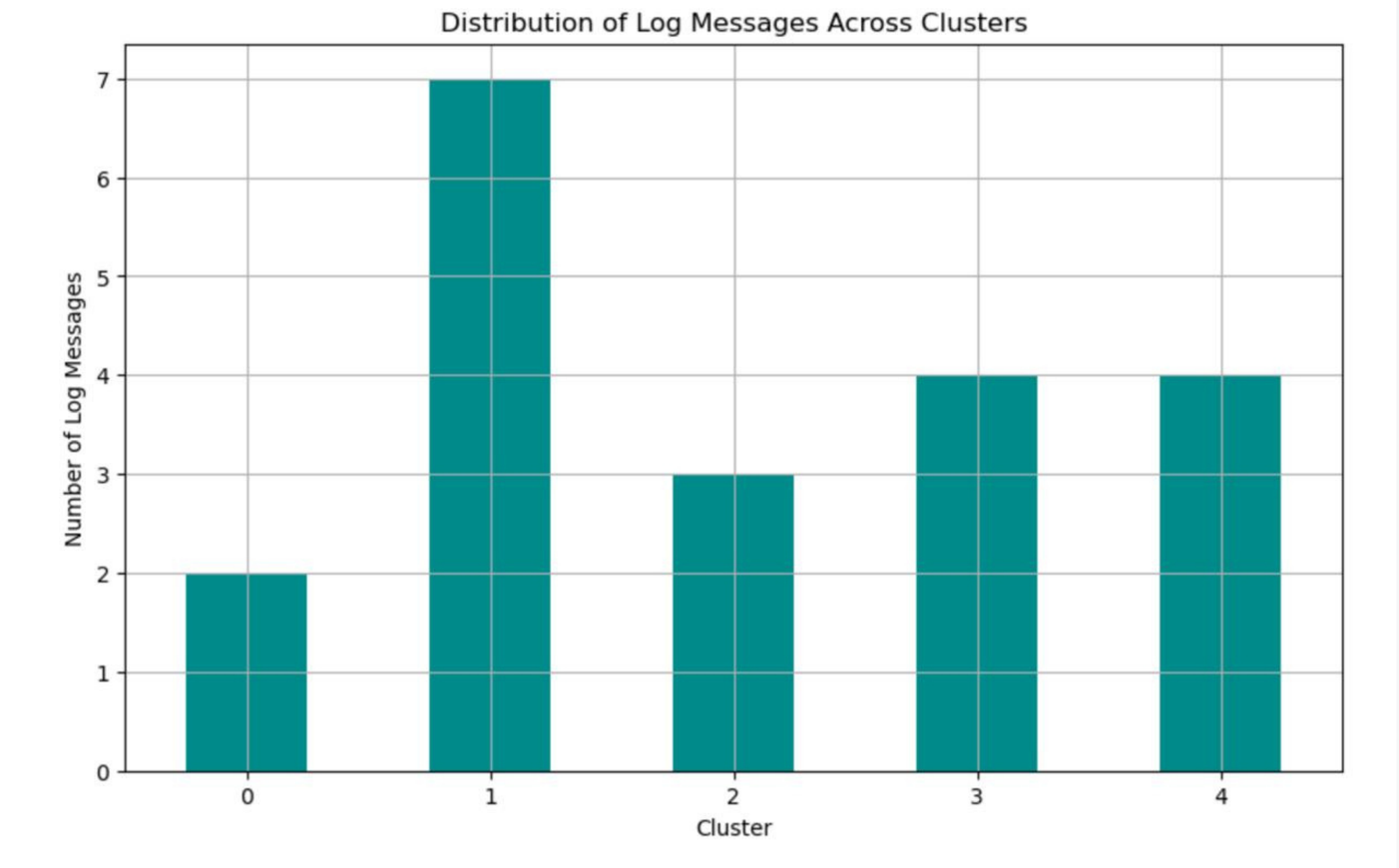
- Uploading user-given log files directly into the **chatbot**.
- Generate **insights** from the data we provide.
- Custom training the Llama 2 language model with our own data to fit our requirements & use cases.

LIMITATIONS

- Lack of data so few anomalies may get undetected
- System Hardware limitations made data handling harder
- The model focuses only on system log files.

INSIGHTS


Process Activity Over Time (Hourly)
Frequency of Process
Hour of the Day
0
200
400
600
800
1000
0
5
10
15
20
Process
kernel
logrotate
cups
named[2288]
named[2275]
named[2305]
named[2281]
named[2306]
syslogd 1.4.1
sendmail




```
Cluster 0: Includes log messages related to page faults and memory allocations, which could indicate issues related to memory management
or hardware errors.
Cluster 1: Consists of messages related to session management, such as opening and closing user sessions. These are routine operations.
Cluster 2: Contains multiple instances of authentication failures, suggesting common security-related incidents or configuration issues
across systems.
Cluster 3: Single entry about user checks passing but the user being unknown, indicating frequent checks for user validation.
Cluster 4: Involves an alert about an abnormal exit, possibly related to errors or unexpected terminations in processes.
```

CHATBOT OUTPUT

```
response = retrieval_chain.invoke({"input": "Derive insights into the following kernel log:[2024-04-20T12:30:15] [error] [kernel]
print(response["answer"])
```

```
Based on the provided kernel logs, here are some insights that can be derived:

1. Memory usage is consistently high: In all the log messages provided, the process is using a significant amount of memory (ab
ove 15000 MB in one case). This could indicate a potential issue with the system's memory management or a misconfiguration in t
he kernel.
2. CPU usage is also high: In several cases, the CPU usage of the process is consistently high ( above 90% in one case), which
could indicate a potential bottleneck in the system or a resource-intensive process.
3. Anomalies are frequent: The log messages suggest that anomalies are occurring frequently, which could indicate a systemic is
sue with the kernel or a problem with the process itself.
4. Out of memory error: The last log message [error] [kernel] [pid:1234] [task:foo] Oops: Kernel panic - not syncing: Out of me
mory indicates that the system has run out of memory, which could be caused by a variety of factors such as a high memory usage
or a lack of memory resources.
5. Potential for data loss or corruption: The frequent anomalies and high memory usage suggest that there is a potential for da
ta loss or corruption if the issue is not addressed promptly.
6. Need for further investigation: Based on the log messages provided, it appears that there is a systemic issue with the kerne
l or a resource-intensive process that needs to be investigated further. This could involve analyzing the process's memory usag
e and CPU usage patterns, as well as checking for any other potential issues in the system.

In conclusion, based on the provided log messages, it appears that there is a potential issue with the system's memory manageme
nt or a resource-intensive process that needs to be addressed promptly to prevent data loss or corruption.
```

THANK YOU

STEALTH IS NOT ABOUT HIDING, IT IS ABOUT BEING UNSEEN UNTIL IT'S YOUR TIME TO MAKE YOUR MOVE