

Performance Analysis Report

Overview

This report presents an analysis of 7 days of app logs, including key performance metrics, anomaly detection, and usage trends.

Codes:

- 200 (OK): The request was successful.
- 400 (Bad Request): The server could not understand the request due to invalid syntax.
- 500 (Internal Server Error): The server encountered an error and could not process the request.
- 504 (Gateway Timeout): The server, acting as a gateway or proxy, did not receive a timely response from the upstream server.

Metrics Summary

	endpoint	avg_response_time	p95_response_time	error_rate	%_slow_requests
0	/api/chat	445.026515	235.65	0.058712	0.009470
1	/api/recommendations	348.311871	155.20	0.066398	0.008048
2	/api/search	471.608081	157.90	0.068687	0.012121
3	/api/user/profile	410.600806	160.00	0.048387	0.010081

Anomaly Detection

- Response Time Anomalies**
 - Number of anomalies: 20
 - These anomalies indicate unusually high or low response times, potentially due to system bottlenecks or transient issues.
- Error Rate Anomalies**
 - No significant anomalies were detected in error rates.
- Endpoint Availability Issues**
 - No endpoints were unavailable during any hour in the dataset.

Insights

- Performance Concerns:** The /api/search endpoint exhibits the highest average response time (471.61 ms) and error rate (6.87%).
- Usage Trends:** Traffic peaks are same for all the 7days.

- **Slow Requests:** The proportion of requests exceeding 1 second is low (<1.3%) across all endpoints, but /api/search has the highest percentage.

Recommendations

1. Optimization

- Investigate the high response times for /api/search and optimize its performance.
- Focus on reducing error rates for /api/recommendations and /api/search.

2. Scalability

- Prepare for scaling resources during peak usage periods identified in the dataset.

3. Anomaly Response

- Automate monitoring of response time anomalies for real-time alerts.