Beyond SLOs

Proactive Failure Prediction Through Anomaly Detection

Gaurav Naresh Mittal

Overview

- The SRE Challenge
- AI in SRE: What and How
- Anomaly Detection Fundamentals
- Case Studies
- Tools and Technologies
- The Future of AI in SRE
- Recommendations and Conclusion

The SRE Challenge

- Key SRE metrics: availability, latency, throughput, error rates
- Monitoring explosion: signal overload creates too much noise
- Alert fatigue and its consequences
- Shift from reactive to predictive reliability

AI in SRE: What

- Intelligent alerting: Filter out the noise and surface only the critical alerts
- Predictive Maintenance: Forecasting potential issues before they impact users.
- Automated Remediation: Triggering corrective actions automatically.
- Capacity Planning: Optimizing resource allocation based on predicted demand.
- Root Cause Analysis: Assisting in quickly identifying the source of problems.
- Incident Management: Automating tasks like ticket creation and prioritization.
- ChatOps Enhancement: Using NLP to interact with systems and run diagnostics via chat.

AI in SRE: How (Part 1)

Types of machine learning applicable to SRE

- Supervised learning (eg. classification of known issues)
- Unsupervised learning (eg. anomaly detection for early warning)
- Reinforcement learning (eg. automated remediation)

Core anomaly detection techniques

- Statistical methods (z-scores, MAD)
- Clustering (k-means, DBSCAN)

AI in SRE: How (Part 2)

AIOps Integration Approaches

- Overlay Model (AI alongside existing tools)
- Embedded Model (AI built into monitoring platforms)
- Automated Remediation (AI driving corrective actions)

Implementation Challenges

- Data quality and quantity requirements
- Training and tuning complexities

Anomaly Detection Fundamental

Anomaly: Deviation from expected system behavior (learned from historical data).

- Point Anomaly: Single, unusual data point.
- Contextual Anomaly: Unusual data point given the context.
- Collective Anomaly: Series of data points that together indicate a problem.

Methods of Detection

- Statistical Methods: Moving averages, exponential smoothing, etc.
- Machine Learning Methods: Clustering, time series forecasting, isolation forest, etc.
- Challenges: Seasonality, noisy data, concept drift, labeled data.

Case Studies

- Netflix: Anomaly detection prevents streaming outages
- LinkedIn: Correlation engines reduce MTTR
- Uber: Predictive scaling with machine learning

Tools and Technologies

Key Open-Source Tools

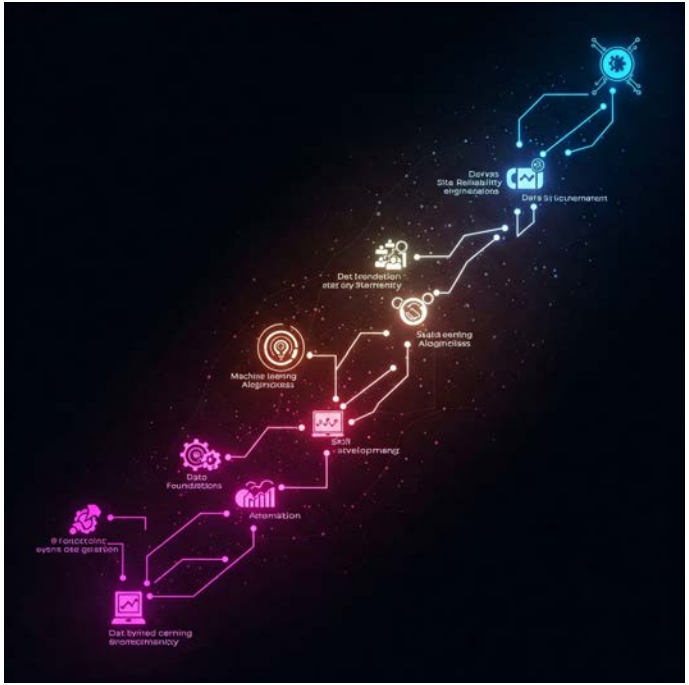
- Prometheus: Metrics collection and storage.
- Grafana: Visualization and dashboarding.
- ELK Stack: Log management and analysis.
- Prophet: Time series forecasting library.
- Scikit-learn: Machine learning library.
- TensorFlow/PyTorch: Deep learning frameworks.

Commercial Platforms: Datadog, New Relic, Dynatrace, Splunk, Amazon CloudWatch, Google Cloud Operations.

The Future of AI in SRE

- Increased Automation: Greater incident response and remediation automation.
- Self-Healing Systems: Auto-detect, diagnose, and fix problems without human intervention.
- More Sophisticated Models: Deep learning for complex systems and data.
- Explainable AI (XAI): Transparent and understandable AI models.
- AI-Driven Observability: Holistic, AI-powered system understanding.

Recommendations: Get Started Now



Focus on specific, high-value use cases.

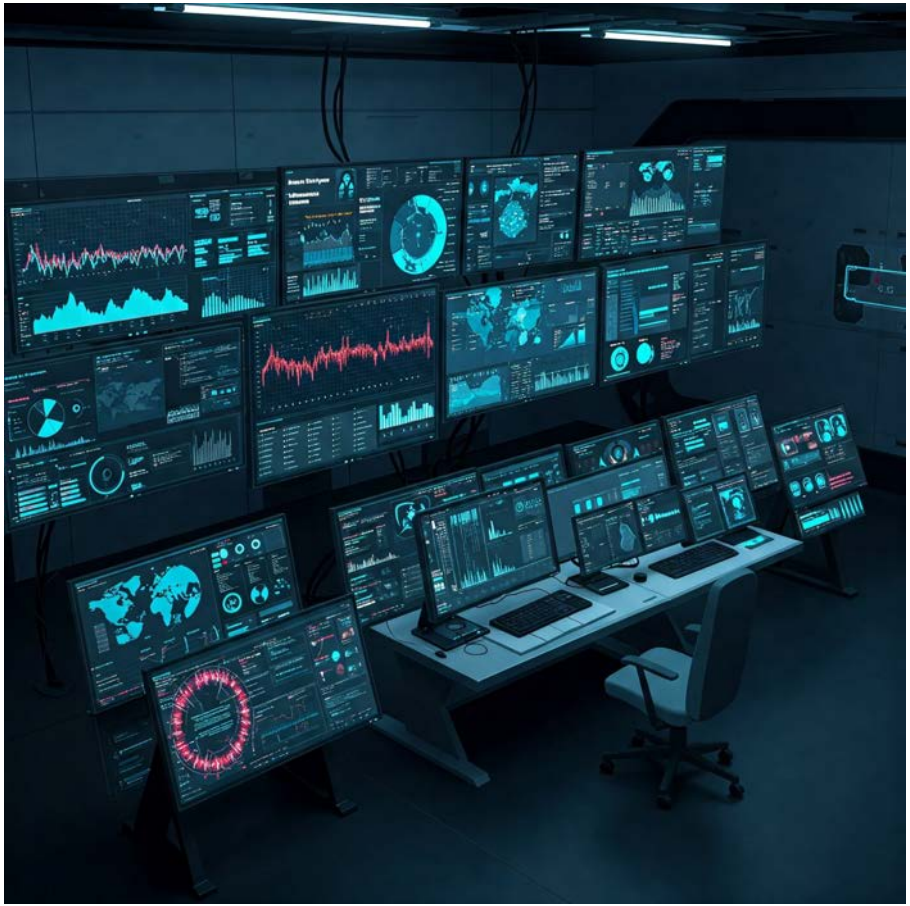
Build the right data foundation.

Start small and follow a roadmap.

Invest in skills development.

Conclusion & Q&A

- Recap: Reduced MTTR, improved reliability, and increased efficiency.
- Call to Action: Explore AI and anomaly detection for SRE
- Start small, focus on a specific use case
- Iterate from there



Thank You