**Name:** Aliyan Ahmed Cheema

**Reg #** FA22-BCE-028

**CyberSentinel Summary**

The research paper titled *"CyberSentinel: An Emergent Threat Detection System for AI Security"* by Dr. Krti Tallam presents a novel single-agent cybersecurity framework designed to proactively detect and respond to emerging threats in AI-driven systems. As artificial intelligence technologies become more integrated into critical infrastructures, the complexity and unpredictability of their behavior expose new vulnerabilities. Traditional security systems, which rely on signature-based detection or manual updates, are often reactive and insufficient against novel attack strategies such as model drift exploitation, prompt injection, and deepfake-based phishing. These threats do not follow known patterns and can bypass conventional defenses, underscoring the urgent need for adaptive and intelligent security solutions.

CyberSentinel addresses this gap by integrating three key components into a unified agent: brute-force attack detection through SSH log analysis, phishing detection using both domain blacklists and heuristic URL scoring, and an Emergent Threat Detector (ETD) that employs unsupervised machine learning models to identify anomalous behavior. The ETD component is particularly innovative as it continuously learns from system activity and adapts to new patterns over time without requiring labeled data. Algorithms such as Isolation Forest and Mahalanobis Distance enable the detection of zero-day attacks, insider threats, and subtle behavioral deviations.

The system is implemented in Python with support for multi-threading and containerized deployment through Docker. It is scalable, with benchmarks showing it can process over a million SSH logs in under three minutes and perform phishing checks at a rate of 10,000 URLs per second. The framework also supports integration with SIEM tools, Slack alerts, and firewall rules for automated mitigation, making it suitable for real-world enterprise deployment.

CyberSentinel contributes to the field of cybersecurity by offering a flexible, explainable, and lightweight solution capable of responding to evolving adversarial tactics. The paper highlights the importance of proactive anomaly detection over reactive rule-matching and proposes future extensions including deepfake detection, adversarial robustness testing, and integration with federated learning and AI safety frameworks. Overall, CyberSentinel represents a significant advancement toward building resilient, AI-aligned defense systems for modern cybersecurity environments.