Tuesday, June 25, 2024

12:10 PM - 12:30 PM

Enhancing Cybersecurity with Cloudera: A Generative AI Approach for the Department of Defense

Tej Tenmattam

Principal Solutions Engineer

Cloudera Government Solutions

Abstract:

The Department of Defense (DoD) is at the forefront of an AI revolution, seeking to harness the power of artificial intelligence (AI), machine learning (ML), and Generative AI (GenAI) to transform its operations. A critical challenge in this endeavor is the need for high-quality data to train language learning models, particularly in the context of cybersecurity. The Defense Information Systems Agency (DISA) faces a pressing issue with the vast amount of cyber data it ingests daily, which requires effective labeling to train models capable of identifying known attacks and detecting potential new zero-day threats.

Cloudera offers a comprehensive solution to this challenge, providing a platform that enables the DoD to establish a continuous data pipeline across global and hybrid data sources. By leveraging Cloudera's advanced data management and analytics capabilities, the DoD can accelerate data access and ensure the quality of data feeding into AI models. This is crucial for deploying future-ready AI programs that can effectively analyze cyber log data, enhancing the capabilities of DISA's analysts in defending the DoD Information Network (DoDIN).

Cloudera's platform supports the development and deployment of Generative AI/LLMs, which can automate the process of data labeling and generate synthetic data for training models. This reduces the manual effort required and increases the efficiency of training robust language learning models. Furthermore, Cloudera's commitment to security ensures that sensitive data is managed and analyzed with the highest level of protection, maintaining the integrity of DoD's cybersecurity efforts.

In this Tech Talk, we will explore how Cloudera's solution can be leveraged to address the DoD's challenges in utilizing cyber data for training language learning models. We will discuss the importance of a continuous data pipeline, the role of Generative AI in augmenting analyst capabilities, and the considerations for deploying AI programs that are future-ready and effective in detecting and defending against cyber threats.