Abhik Risal

30419133

Week 11 Lab Activities – Exploring Emerging Topics in Cybersecurity

Topic: AI and Adversarial Cybersecurity

Reflection

Artificial Intelligence (AI) is rapidly transforming cybersecurity, strengthening both defensive and offensive capabilities. Modern Security Operations Centers (SOCs) now rely on AI-driven analytics to detect anomalies, predict attack patterns, and automate responses. However, the same technology empowering defenders is also being weaponized by adversaries, creating a new category of risk—**adversarial cybersecurity.** Attackers now use AI to automate phishing campaigns, generate deepfake videos, bypass CAPTCHAs, and evade traditional detection models.

One striking example is the 2020 deepfake-based scam in which criminals used an AI-cloned CEO voice to trick a financial officer into transferring $243,000. Similarly, AI-generated phishing emails have become nearly indistinguishable from legitimate corporate communication. These cases reveal how AI lowers the barrier to sophisticated attacks while blurring the line between authentic and synthetic data.

On the defensive side, AI-powered tools like **Darktrace**, **CrowdStrike Falcon**, and **Google Chronicle** are improving real-time threat visibility by learning network behavior and detecting deviations. Yet, these systems are also vulnerable to **adversarial machine learning**—where attackers poison datasets or manipulate inputs to cause false negatives. The challenge is ensuring transparency, explainability, and bias-free decision-making in AI models.

This topic is deeply relevant to my future in cybersecurity. As AI-driven defense becomes standard, professionals must understand both its potential and limitations. The future SOC analyst will not only interpret AI alerts but also evaluate the trustworthiness of algorithms themselves. This dual awareness—of how AI defends and how it can be exploited—will define the next generation of cybersecurity expertise.
