Will Young
Student ID: 924230057
Week 1 Report
1/30/2026

Background of research project: Working at a regional managed service provider (MSP) as a network engineer, I support many clients across different environments. A subset of this work is performing vulnerability and penetration tests, where the results tend to be lengthy reports explaining what was found. Translating these reports into concrete, actionable, environment-specific hardening steps typically requires multiple tiers of engineers. Typically, our senior staff interpret the findings, then they familiarize themselves with the client's unique tech stack and design technical and administrative controls to address deficiencies. This step is the most time-consuming as it requires expertise and years of experience across different domains to develop this remediation roadmap.

These senior engineers are also responsible for various complex, high-value tasks across our client base, making their time already scarce and expensive. This challenge is not unique to my company but rather an industry wide bottleneck. My solution is to leverage generative AI to resolve this problem. At a high level, we ingest unstructured data, whether it is a vulnerability report, network telemetry or traffic from an incident, then we map it to MITRE ATT&CK techniques and summarize the event in natural language, ending in a set of concrete remediation steps customized to the clients tech stack that a level 1 or level 2 technician can implement. The goal is to free up the need for senior level engineers while keeping humans in the loop to implement the final recommended policy.

What it is: Now that I've discussed the background of this project, I will dive into the details a bit more. I will still be a little high level, as some details may dynamically change as we progress through the semester. I will break this section up into 4 sub sections and discuss each section shown in the image below.

Section 1.) This section will be the ingestion pipeline for the project. It will take in all sorts of structured and unstructured data like vulnerability report PDF's, Excel spreadsheets, CSV logs and PCAP files and normalize them for structured summaries with important raw excerpts. This will focus the model by cutting down on the noisy, unuseful data while preserving key evidence and patterns. This process will involve special python libraries to extract key information and create uniformed JSON packets to be shipped to AI for further correlation.

Section 2.) This section will build off of the work in Esposito's thesis paper you sent towards the end of last semester. We will incorporate the Reporter and Mapper architecture, where each role has a specialized model. They used GPT-4o for the reporter

role and fine-tuned Mistral-7B, using LoRA, for the Mapper. I will spend some considerable time to look at hugging face for more modern models, possibly using Gemini for the reporter role. The Reporter role would take the normalized JSON packet and construct a single technical sentence to summarize the event. This is then fed into the Mapper model, which analyzes that sentence, and correlates it to a MITRE ATT&CK ID. These results are then fed into our correlation agent in section 3.

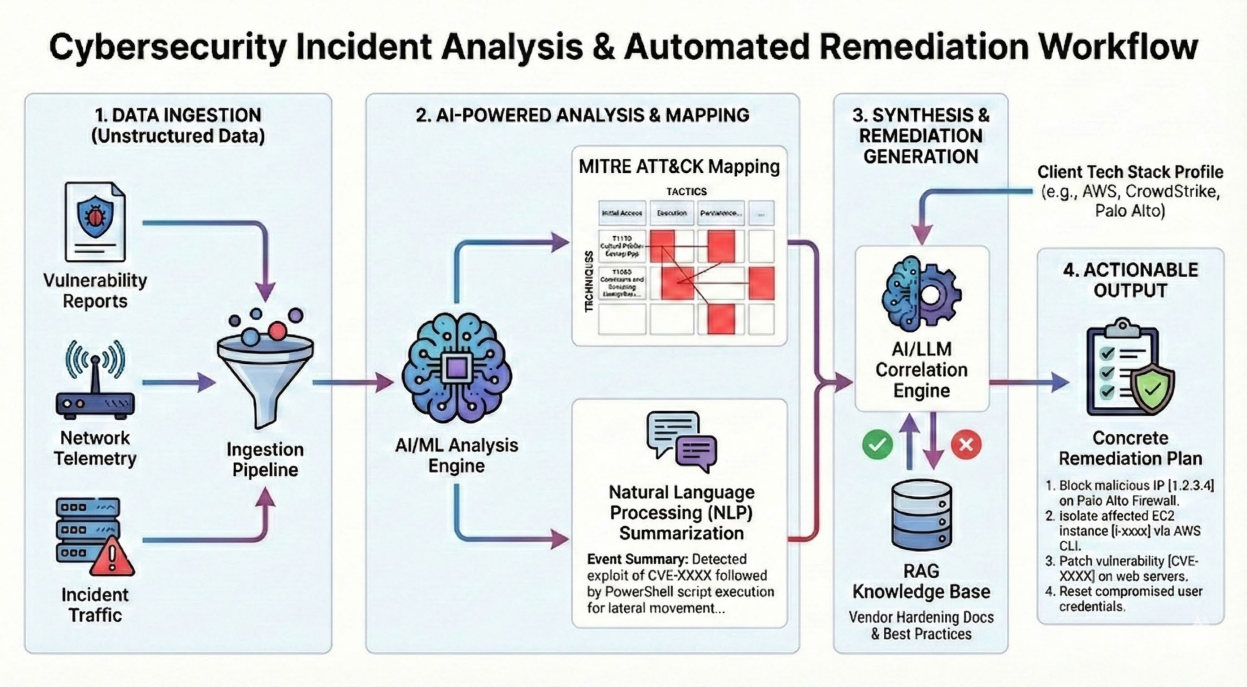

Figure 1: Project Workflow

Section 3.) This section takes in the ATT&CK IDs and the Reporter's sentence from the previous section and adds client environment vendor context. This is crucial for developing a custom remediation plan for the client's tech stack where they implement their technical and administrative controls securing their digital footprint. For example, we would say client X uses Meraki MX, MS and MR, Active Directory, Windows Server 2019, Windows endpoints with Datto EDR/AV and O365. With this additional context, and the information provided from section 2, our model will reach out to its RAG database to look for vendor appropriate hardening best practices. This database will be populated with documentation from popular enterprise vendors as well as CIS benchmarks and ATT&CK mitigation documentation. The LLM will then build a list of vendor-tailored remediation steps. We will also implement a Self-RAG check to weed out any hallucinations, ensuring we are providing real guidance based on real events.

Section 4.) This section will take the remediation steps from section 3 and extrapolate the remediation steps for a L1 or L2 engineer. The output might include a step-by-step or click-by-click walk through of the vendors UI in order to get to the appropriate screen where these technical changes need to be made. This may look like “Meraki dashboard -> Security & SD-WAN -> Threat Protection ->Enable Advanced Malware Protection & Enable Intrusion Detection and Prevention.” The most important aspect of this step is it is the human in the loop portion. Security in general has very little room for inaccuracy, and LLM’s do confidently produce made up content. Even with some of the mitigation steps we will be implementing, it is important that a human can look at final product, see what has happened and review the proposed remediation step to determine if; yes that is correct and I will move forward or if it is incorrect and may need additional work to get to the right solution.

Why it is important: This project is incredibly important because there is a substantial shortage of highly qualified security engineers with enough experience and cross-domain knowledge to look at these technical reports and then implement and execute a plan to remediate each issue that was identified. In large multi-national corporations, there are well-built teams, distributed across time zones and tiers of expertise that deal with these issues, but most companies don’t have that luxury or budget. The shortfall of qualified security engineers in the United States is only growing, while the threats to companies’ digital infrastructure is exponentially expanding. This project aims to help fill this gap and provide lower-level engineers the guidance and confidence to help secure their clients’ digital environment.

Current progress: Currently, I have finalized what I think will be a very in-depth and interesting project. In my professional work, I have used many different security and reporting products, and they all lack this last mile remediation step. This week I plan on creating a general timeline of milestones to help keep on track as there is a lot going into this project. I will be reviewing LLM models on hugging face and reviewing project adjacent research papers as well.