CS486C – Senior Capstone Design in Computer Science

Project Description

Project Title: Leveraging AI to Automate Cybersecurity Testing Process	
Sponsor Information: HIGHVIZ SECURITY Uncovering the Unseen	Rick Belisle, CEO HighViz Security LLC rbelisle@highviz.tech

Project Overview:

HighViz Security is a cybersecurity company that specializes in conducting penetration testing and application security assessments. In essence, we are paid ethical hackers, hired by corporations to find and exploit existing vulnerabilities within their enterprise networks with the goal that we identify and help them fix their security problems before malicious individuals attack them. The team at HighViz has specialized in this area for over 20 years, and our niche in this space is focused on highly customized manual testing techniques. This process, however, takes a significant amount of time to do well. Unlike real hackers, when it comes to cybersecurity testing, we are at a disadvantage in that we have a limited amount of time to conduct our testing so, to be successful, we must be able to provide full testing coverage as quickly and efficiently as possible.

Our goal for this project is to ultimately shorten the first phase of testing, which typically relies on automated vulnerability scanning software that can quickly identify the low hanging fruit and help understand the types of systems and services that are available for further testing. The problem is that these tools typically generate a lot of false positives as well as creating a lot of data that needs to be evaluated before continuing the testing process. When scanning large enterprise networks with thousands of systems the scanning logs alone can quickly become unmanageable.

Ultimately, with the help of our Capstone team, we would like to develop an automated process that can evaluate the data generated by a commercial scanning tool (Tenable's Nessus Security Scanner) and apply some logic that can make analyzing all this data easier. Some of the key requirements will be to:

- Identify and annotate known or expected false positives
- Adjust risk levels of vulnerabilities based on real world experience
- Group findings that all fall under similar categories (i.e. missing patches)
- Highlight vulnerabilities with known exploits
- Summarize key findings into meaningful data points that are easily digested (total number of high/medium/low findings, total number of available services, total number of vulnerabilities with existing exploits, etc.)
- Any other interesting manipulations that might be possible once the data is better organized

HighViz has a working solution that covers a lot of these areas that leverages SQL to normalize the data and map it to known values from our library of existing findings. This existing process, however, is very ad hoc and is based solely on a static mapping process. This makes it fragile and requires constant time-consuming updating of the mapping logic. A successful Capstone team would take this process to the next level by leveraging the use of AI to help streamline the analysis process. The AI powered process will analyze the data for us, allowing the tester to quickly review the critical information and move to the next phase of testing faster, saving countless wasted hours massaging the data looking for the information they need.

This process would need to work in tandem with a tester and learn based on what the tester ultimately does with the vulnerability data. For example, if a tester indicates vulnerabilities X, Y, and Z were exploited or certain issues were indeed false positives, the AI will need to remember this and apply it to future findings.

The use of AI comes with a caveat, however. HighViz is not willing to trust the security of our clients to any existing commercial AI tool and is ultimately looking to create a local AI assistant that does not send any of the data collected out of the local protected environment. The security of the data in this process is paramount and cannot be out of our direct control or left exposed in a 3rd party solution that could put our client at risk.

In the end, as a consulting organization, our success is solely based on the ability of the tester to get the job done. The faster they can complete the goals of the testing the more projects we can do, and the more successful we become. Successfully completing this project would serve multiple key goals for us to include:

- Decrease overall testing time, or at least, maximize the tester's time conducting advanced testing techniques.
- Help standardize the testing process when testing as a team, ensuring all testers are reporting findings in a consistent way.
- Streamline the reporting process by capturing key findings quickly, and in a standardized format.
- Helping to highlight findings that may not be immediately obvious in the overwhelming amount of data generated by automated tools.
- Ability to expand on the tool in future iterations to incorporate other security scanning software beyond Nessus.

Knowledge, skills, and expertise required for this project:

- Basic file operations
- Programming skills in Python
- Basic SQL Lite operations
- Advanced AI programming skills

Equipment Requirements:

- Any standard computing platform, Mac OS is preferred but not required

Software and other Deliverables:

- Final report detailing the design and implementation of the product in a complete, clear and professional manner. This document should provide detailed usage instructions, data mapping details, detailed AI training process, and a strong basis for how future development of the product can incorporate data from other tools.
- Complete professionally documented codebase delivered both as a repository in GitHub, and as a physical archive on a USB drive.