SHANE REALL, Dept of Science and Technology, Fairmont State University, Fairmont, WV 26554.
Evaluating the Security Awareness and Code Reliability of AI Models Against Common Software Vulnerabilities

For this project we explored how Large Language Models (LLMs) understand and handle software security, with us mostly focusing on Common Weakness Enumeration (CWEs). Today LLMs are commonly used for programming shortcuts and other common programming issues, this has led to many programmers relying on them in a professional setting. This could cause many security flaws which could be exploited. For this I tested 4 LLMs, including their paid versions, to test prompts that could be vulnerable to 18 different CWEs, resulting in 144 experiments. These will be testing CWEs such as Improper Authentication, Integer Overflow, OS Command Injection, and Insufficiently Random Values use, and will be testing them with languages such as C, Java, and Python. The generated code is then tested, analyzed, and categorized based on a rubric. In comparison to previous studies, this study looks at a large set of security weaknesses and comparison in free vs paid account on these LLMs. These results aim to teach programmers and software developers how to integrate AI programming tools correctly, safely, and responsibility with any programming field.