ZACH PFALTZGRAFF, 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

This study evaluates the ability of AI coding assistants to recognize and mitigate common software vulnerabilities defined by the Common Weakness Enumeration (CWE). As AI tools become increasingly integrated into software development, concerns remain regarding their reliability in producing secure code. Four AI models, including both free and paid versions, were assessed using standardized prompts targeting 11 critical CWE categories, which has 88 total test cases. Vulnerabilities examined include Cross-Site Scripting, SQL Injection, Buffer Overflows, Use-After-Free errors, and Hard-Coded Credentials across multiple programming languages.

Generated code outputs were analyzed and classified as secure, partially secure, or insecure based on established secure coding practices. Results indicate variability in model performance, with some models consistently producing secure implementations while others frequently introduced or failed to mitigate known vulnerabilities. Statistical comparisons highlight differences in reliability across both model type and vulnerability category.

These findings demonstrate that while AI coding assistants can improve development efficiency, they do not consistently ensure secure code generation. Developers should exercise caution and incorporate additional validation measures when using AI tools, particularly in security-sensitive domains. This research provides practical insights into the limitations of AI-assisted coding and supports the need for improved model training focused on secure software development.