

DeepSeek R1, Llama3, and Codestral Vulnerability Detection Across Cyclomatic Complexity Levels

Assignee Research

June 4, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does the vulnerability detection performance of Deepseek R1 on the Big-Vul dataset vary across different levels of cyclomatic complexity compared to Llama3 and Codestral. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: LLMs in Code Vulnerability Analysis: A Proof of Concept. Research question: How does the vulnerability detection performance of Deepseek R1 on the Big-Vul dataset vary across different levels of cyclomatic complexity compared to Llama3 and Codestral?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.8/10.

3 Results

15 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 4.8/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

References

- <http://arxiv.org/abs/2503.10486v2>
- <http://arxiv.org/abs/2601.08691v1>
- <http://arxiv.org/abs/2512.06483v1>