

Multimodal vs. Text-Only Llama3 Variants in OWASP Top 10 Vulnerability Detection

Assignee Research

June 4, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: How do multimodal Llama3 variants compare to text-only variants in detecting OWASP Top 10 vulnerabilities when evaluating response safety metrics like accuracy and precision under adversarial. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Benchmarking LLAMA Model Security Against OWASP Top 10 For LLM Applications. Research question: How do multimodal Llama3 variants compare to text-only variants in detecting OWASP Top 10 vulnerabilities when evaluating response safety metrics like accuracy and precision under adversarial perturbations?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.7/10.

3 Results

14 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.7/10.

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/2601.19970v1>
- <http://arxiv.org/abs/2507.22398v3>
- <http://arxiv.org/abs/2407.13111v1>