Stratified Sampling Enhances Robustness in Llama3 and Codestral for Cross-Domain Code Vulnerability Detection

Assignee Research

June 2, 2026

Abstract

This report synthesises findings from 9 peer-reviewed papers addressing the following research question: Does the use of stratified sampling improve the robustness of Llama3 and Codestral in cross-domain code vulnerability detection, as evaluated by accuracy on a mixed-domain benchmark like MBXD. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Demystifying Instruction Mixing for Fine-tuning Large Language Models. Research question: Does the use of stratified sampling improve the robustness of Llama3 and Codestral in cross-domain code vulnerability detection, as evaluated by accuracy on a mixed-domain benchmark like MBXD?.

2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.8/10.

3 Results

9 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 6.8/10.

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/2306.08568v2
- http://arxiv.org/abs/2503.08977v1
- http://arxiv.org/abs/2312.10793v3