Instruction-Tuning Data Diversity and Cross-Domain Zero-Shot Code Generation in Llama-2 Models

Assignee Research

June 3, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: How does the diversity of instruction-tuning data affect the cross-domain zero-shot code generation capability of Llama-2 models, as measured by pass@1 accuracy on HumanEval across Python,. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: AeroGen: Enhancing Remote Sensing Object Detection with Diffusion-Driven Data Generation. Research question: How does the diversity of instruction-tuning data affect the cross-domain zero-shot code generation capability of Llama-2 models, as measured by pass@1 accuracy on HumanEval across Python, JavaScript, and Java domains?.

2 Methodology

Systematic literature search across multiple databases yielded 4 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.8/10.

3 Results

4 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.8/10.

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/2411.15497v3
- http://arxiv.org/abs/2410.12381v3
- http://arxiv.org/abs/2510.21391v1