Robustness Disparities of Qwen3 and Qwen2-1.5B Against Adversarial Docstring Perturbations in Low- and High-Resource Programming Languages

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does the robustness of Qwen3 and Qwen2-1.5B against adversarial docstring perturbations vary when evaluated on the HumanEval-X dataset for low-resource programming languages compared to. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Evaluating the Limits of Large Language Models in Multilingual Legal Reasoning. Research question: How does the robustness of Qwen3 and Qwen2-1.5B against adversarial docstring perturbations vary when evaluated on the HumanEval-X dataset for low-resource programming languages compared to high-resource ones?.

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 3.3/10.

3 Results

15 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.3/10.

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/2509.22472v1
- http://arxiv.org/abs/2408.04430v3
- http://arxiv.org/abs/2308.09895v6