Failure Rates of Gemini-2.5-Flash on Edge-Case Coding Tasks Under Varying Input Conditions

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: What is the impact of input length and complexity on the failure rate of Gemini-2.5-Flash for edge-case coding tasks, and how does it scale with model size or computational resources. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Precise Length Control in Large Language Models. Research question: What is the impact of input length and complexity on the failure rate of Gemini-2.5-Flash for edge-case coding tasks, and how does it scale with model size or computational resources?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.3/10.

3 Results

16 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 4.3/10.

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/2507.12472v1
- http://arxiv.org/abs/2412.11937v1
- http://arxiv.org/abs/2303.12869v1