

Summary Length and False Positive Rates in Reference-Free Factuality Metrics for Adversarial Long Documents

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: What is the correlation between summary length and the false positive rate of reference-free factuality metrics when evaluating sliding window attention outputs on adversarially perturbed long. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Stress Testing Factual Consistency Metrics for Long-Document Summarization. Research question: What is the correlation between summary length and the false positive rate of reference-free factuality metrics when evaluating sliding window attention outputs on adversarially perturbed long documents?.

2 Methodology

Systematic literature search across multiple databases yielded 10 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.2/10.

3 Results

10 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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/2309.12455v2>
- <http://arxiv.org/abs/2511.07689v2>
- <http://arxiv.org/abs/2406.19276v1>