

Token Misalignment Threshold Effects on Hallucination Rates in Vicuna-13B and Baichuan-2

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does varying the TAE token misalignment threshold affect the hallucination rates of Vicuna-13B and Baichuan-2 across different domains in the FactCC and HalluEval benchmarks. Large language models (LLMs) have demonstrated remarkable performance on a variety of natural language tasks based on just a few examples of natural language instructions, reducing the need for extensive feature engineering. However, most powerful LLMs are closed-source or 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Baichuan 2: Open Large-scale Language Models. Research question: How does varying the TAE token misalignment threshold affect the hallucination rates of Vicuna-13B and Baichuan-2 across different domains in the FactCC and HalluEval benchmarks?.

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 6.5/10.

3 Results

15 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 6.5/10.

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/2507.20836v4>
- <http://arxiv.org/abs/2309.10305v4>
- <http://arxiv.org/abs/2310.00741v2>