Retrieval-Augmented 7B Models Match 70B Accuracy on Specialized Religious Datasets

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: Can a 7B model with retrieval-augmentation achieve comparable accuracy to a 70B model on specialized religious datasets while maintaining a response latency under 500ms per query in a batch size of 1. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: MUST-RAG: MUSical Text Question Answering with Retrieval Augmented Generation. Research question: Can a 7B model with retrieval-augmentation achieve comparable accuracy to a 70B model on specialized religious datasets while maintaining a response latency under 500ms per query in a batch size of 1?.

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 4.2/10.

3 Results

16 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 4.2/10.

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/2507.23334v2
- http://arxiv.org/abs/2503.16581v1
- http://arxiv.org/abs/2404.14464v1