Adversarial Robustness of MA-DPR-Based RAG Systems vs. Lexical Retrieval on AdversarialQA

Assignee Research

June 2, 2026

Abstract

This report synthesises findings from 3 peer-reviewed papers addressing the following research question: How does the performance of MA-DPR-based RAG systems degrade under adversarial attacks compared to lexical retrieval methods when evaluated on the AdversarialQA benchmark for robustness. Abstract Transformer-based language models (TLMs) have widely been recognized to be a cutting-edge technology for the successful development of deep-learning-based solutions to problems and applications that require natural language processing and understanding. Like for other. 5 claims were extracted from source literature; 5 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 8.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Bringing order into the realm of Transformer-based language models for artificial intelligence and law. Research question: How does the performance of MA-DPR-based RAG systems degrade under adversarial attacks compared to lexical retrieval methods when evaluated on the AdversarialQA benchmark for robustness?.

2 Methodology

Systematic literature search across multiple databases yielded 3 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.0/10.

3 Results

3 papers retrieved. 5 claims extracted; 5 independently verified. Quality review score: 8.0/10.

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.

5 Extracted Claims

Claim	Verified	Confidence
Transformer-based language models (TLMs) are recognized as a cutting-edge technology for developing deep-learning-based	✓	0.37
TLMs have pushed the state-of-the-art of AI approaches for many tasks of interest in the legal domain.	✓	0.32
The first Transformer model was proposed about six years ago.	✓	0.21
BERT and related models represent a major reference in the legal domain.	✓	0.29
This article provides the first systematic overview of TLM-based methods for AI-driven problems and tasks in the legal s	✓	0.37

References

- https://doi.org/10.18653/v1/2022.naacl-main.226
- https://doi.org/10.4230/tgdk.1.1.2
- https://doi.org/10.1007/s10506-023-09374-7