

Spelling-Correction Pre-Processing Impact on Dual-Encoder Retrieval Accuracy in NaturalQuestions

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does integrating spelling-correction pre-processing modules affect the retrieval accuracy of dual-encoder models on the NaturalQuestions benchmark compared to query-type specific retriever. 12 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Analysing the Robustness of Dual Encoders for Dense Retrieval Against Misspellings. Research question: How does integrating spelling-correction pre-processing modules affect the retrieval accuracy of dual-encoder models on the NaturalQuestions benchmark compared to query-type specific retriever ensembles?.

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 2.5/10.

3 Results

15 papers retrieved. 12 claims extracted; 0 independently verified. Quality review score: 2.5/10.

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.

5 Extracted Claims

Claim	Verified	Confidence
On clean questions, data augmentation, contrastive learning, and their combination do not harm the retrieval performance	×	0.12
All robustification approaches (Data augm., CL, Data augm. + CL) perform significantly better than the original Dual Enc	×	0.06
The combined approach of data augmentation and contrastive learning achieves the highest performance among all tested me	×	0.08
Robustness of Dual Encoders deteriorates when typos are restricted to non-stopwords compared to when typos appear random	×	0.09
The most significant performance losses occur when typos appear in discriminative utterances (words overlapping with the	×	0.04
The combined approach of data augmentation and contrastive learning remains the best performing method across all three	×	0.07
There is a strong positive correlation between the frequency of typed words in the training set and retrieval performan	×	0.09
Retrieval performance drops significantly as the frequency of typed words in the training set decreases.	×	0.05
On the Natural Questions test set with random typos, the original Dual Encoder achieves an AR@5 of 49.52.	×	0.05
On the Natural Questions test set with random typos, the Dual Encoder with Data Augmentation and Contrastive Learning ac	×	0.09
On the Natural Questions test set with typos in discriminative utterances, the original Dual Encoder achieves an AR@5 of	×	0.05
On the Natural Questions test set with typos in discriminative utterances, the Dual Encoder with Data Augmentation and C	×	0.09

References

- <http://arxiv.org/abs/2204.00716v2>

- <http://arxiv.org/abs/2205.02303v1>
- <http://arxiv.org/abs/2504.04062v1>