

Sentence-T5 and MPNet Embedding Fusion for Cross-Domain Retrieval on HotpotQA

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: Does combining Sentence-T5 and MPNet embeddings improve cross-domain retrieval accuracy on HotpotQA when models are trained exclusively on TriviaQA. Modern information retrieval (IR) models, trained exclusively on standard <query, passage> pairs, struggle to effectively interpret and follow explicit user instructions. We introduce InF-IR, a large-scale, high-quality training corpus tailored for enhancing retrieval models in. 10 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.4/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Towards Better Instruction Following Retrieval Models. Research question: Does combining Sentence-T5 and MPNet embeddings improve cross-domain retrieval accuracy on HotpotQA when models are trained exclusively on TriviaQA?.

2 Methodology

Systematic literature search across multiple databases yielded 4 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.4/10.

3 Results

4 papers retrieved. 10 claims extracted; 0 independently verified. Quality review score: 2.4/10.

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
The proposed marginal negative sampling strategy reduces the computational complexity of the denominator from combinator	×	0.02
Objective I (Univariate Conditional Modeling) models three conditional distributions: $P(P I, Q)$, $P(I P, Q)$, and $P(IQ P)$	×	0.04
The multivariate objective (Eq. 7) introduces a larger set of hard negatives compared to the univariate contrastive obje	×	0.02
The FollowIR-7B model was trained on 50.7K instruction samples and 5.9M query-passage pairs.	×	0.12
The Promptriever model was trained on 9.9K instruction samples, 9.9K query samples, and 16.1K passage samples.	×	0.07
Training e5-base-v2 with InF-Embed resulted in a performance score of 14.0 on one benchmark compared to 13.4 for the bas	×	0.05
Training e5-large-v2 with InF-Embed resulted in a performance score of 24.3 on one benchmark compared to 23.8 for the ba	×	0.06
Training ModernBERT-base with InF-Embed resulted in a performance score of 10.0 on one benchmark compared to 4.29 for th	×	0.05
The multivariate formulation exhibits greater robustness to competition-related issues compared to the univariate object	×	0.00
The FollowIR dataset contains 104 instruction samples.	×	0.04

References

- <http://arxiv.org/abs/2505.21439v1>

- <http://arxiv.org/abs/1811.08772v1>
- <http://arxiv.org/abs/2007.00166v1>