Scaling Parameter Effects in Contrastive Pre-training and Alignment Fine-tuning for Zero-Shot Dense Encoder Retrieval

Assignee Research

June 11, 2026

Abstract

Dense retrievers utilize pre-trained backbone language models (e.g., BERT, LLaMA) that are fine-tuned via contrastive learning to perform the task of encoding text into sense representations that can be then compared via a shallow similarity operation, e.g. inner product. Recent research has questioned the role of fine-tuning vs. that of pre-training within dense retrievers, specifically arguing that retrieval knowledge is primarily gained during pre-training, meaning knowledge not acquired during pre-training cannot be sub-sequentially acquired via fine-tuning. We revisit this idea here as th

1 Introduction

This paper examines: Pre-training vs. Fine-tuning: A Reproducibility Study on Dense Retrieval Knowledge Acquisition. Research question: What is the comparative impact of scaling model parameters during contrastive pre-training versus alignment fine-tuning on the zero-shot retrieval performance of dense encoders across heterogeneous domains like PubMedQA and ScienceQA?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.7/10.

3 Results

11 papers retrieved. 5 claims extracted; 5 independently verified. Quality review score: 8.7/10.

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 study investigates the consistency of embedding knowledge between dense retrievers and their untrained backbone coun	✓	0.19
Fine-tuned dense retrieval models encode knowledge in a manner consistent with their pretraining initialization.	✓	0.24
Linear probing involves training a simple linear classifier on embeddings extracted from different layers of a model.	✓	0.27
The task considered by Reichman and Heck is that of distinguishing relevant and irrelevant documents with respect to a q	✓	0.23
If a linear classifier achieves high accuracy, it suggests that the embeddings contain structure.	✓	0.19

References

- http://arxiv.org/abs/2405.19204v1
- http://arxiv.org/abs/2307.00589v2
- http://arxiv.org/abs/2505.07166v1