中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES

Reproducibility Analysis and Enhancements for Multi-Aspect Dense Retriever with Aspect Learning

Keping Bi, Xiaojie Sun, Jiafeng Guo, Xueqi Cheng
Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS)
CAS Key Lab of Web Data Science and Technology

ECIR 2024 GLASGOW

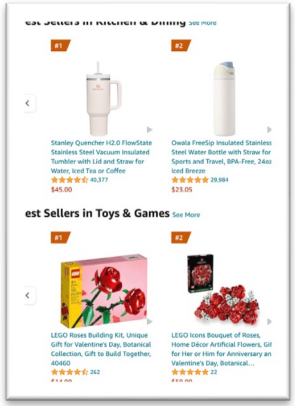
Introduction

What is multi-aspect dense retrieval? Why is it important?

Multi-aspect Dense Retrieval

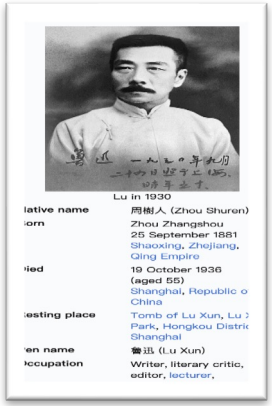
- In structured item retrieval, aspect information is critical.

Product Retrieval



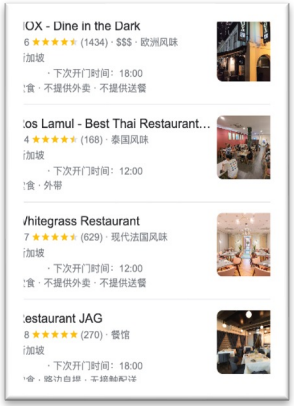
(brand, category, color)

People Retrieval



(nationality, occupation)

Restaurant Retrieval



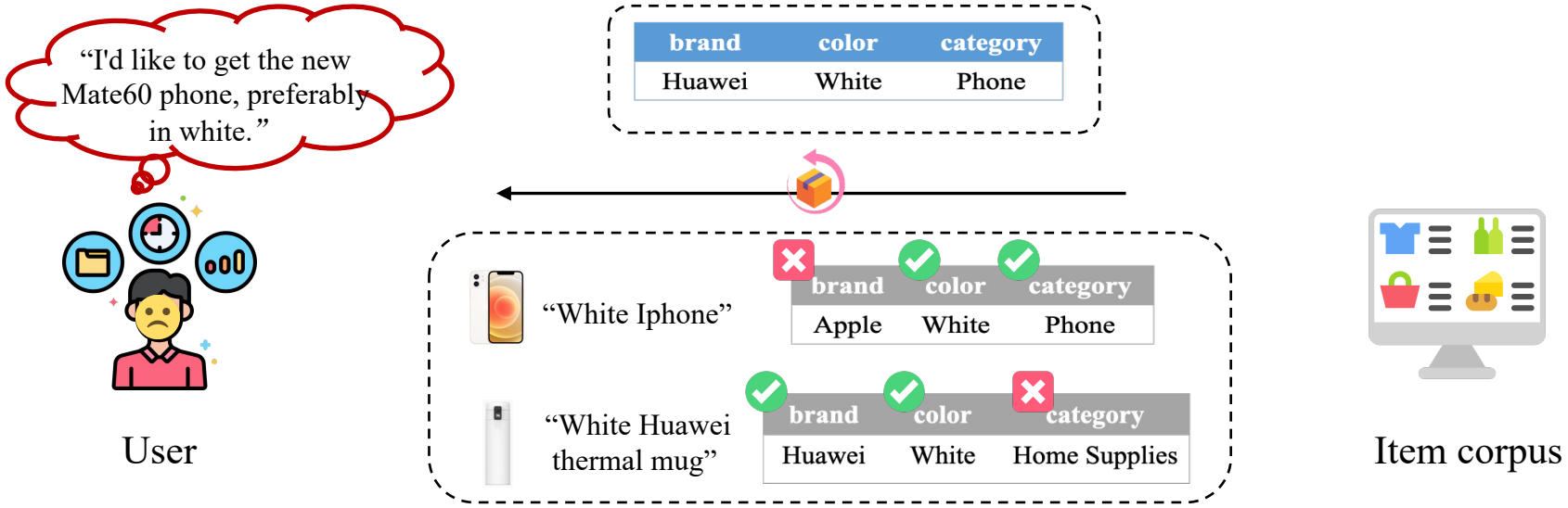
(cuisine, rating)

Clothes Retrieval

(texture, category)

Multi-aspect Dense Retrieval

- Multi-aspect dense retrieval aims to incorporate aspect information into dual encoders to facilitate relevance matching.



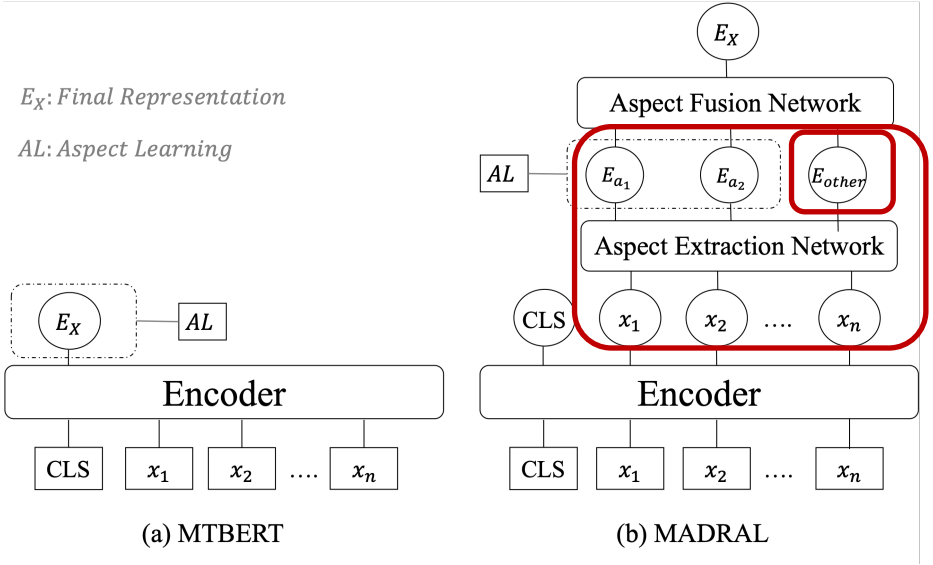
Background

Current challenges in this field

Multi-aspect Dense Retrieval

Three major components:

- **Aspect Extraction**: explicit aspects + a special aspect "other" for implicit semantics
- **Aspect Learning**: predict the value IDs of an aspect
- **Aspect Fusion**: produce the final query/item representation



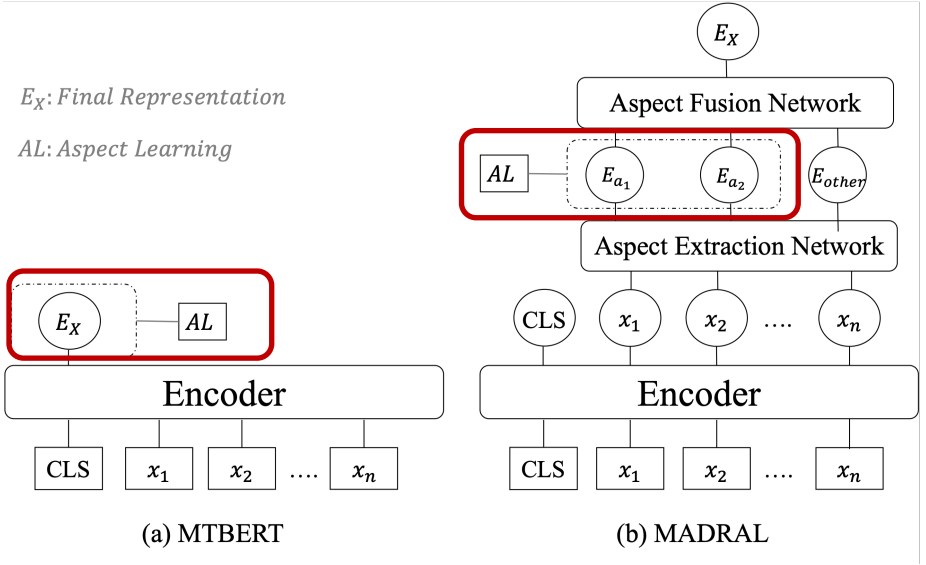
A SOTA Multi-aspect Dense Retriever

Multi-aspect dense retrieval. Kong et al. KDD 2022

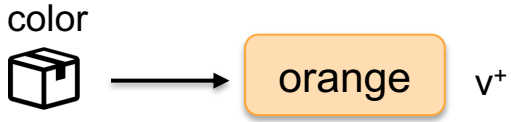
Multi-aspect Dense Retrieval

Three major components:

- **Aspect Extraction**: explicit aspects + a special aspect "other" for implicit semantics
- **Aspect Learning**: predict the value IDs of an aspect
- **Aspect Fusion**: produce the final query/item representation



$$\mathcal{L}^{a}_{AspectPrediction} = -\sum_{v^{+}\in\mathcal{A}_a} \log \frac{\exp(E_a \cdot E_{v^{+}})}{\sum_{v\in V_a}\exp(E_a \cdot E_v)}$$

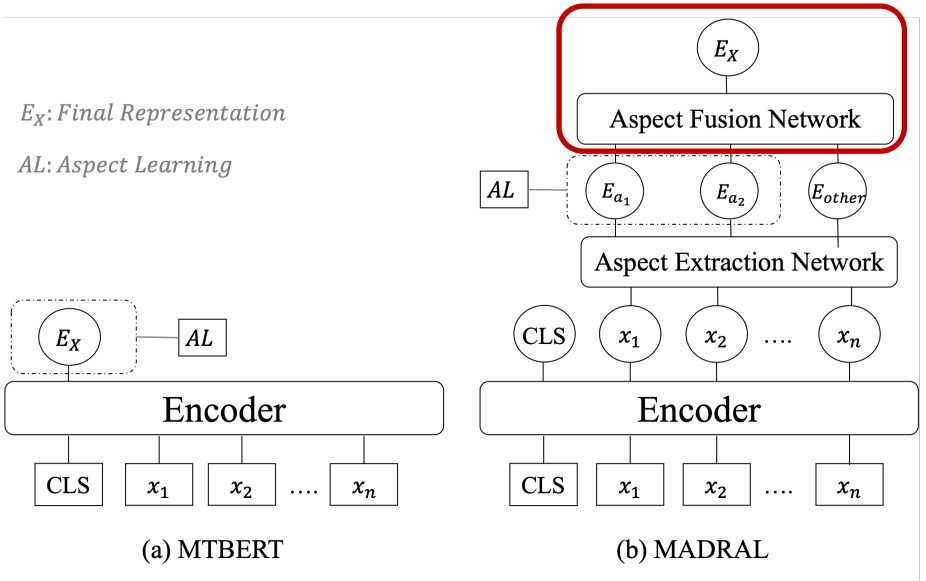


Multi-aspect dense retrieval. Kong et al. KDD 2022

Multi-aspect Dense Retrieval

Three major components:

- **Aspect Extraction**: explicit aspects + a special aspect "other" for implicit semantics
- **Aspect Learning**: predict the value IDs of an aspect
- **Aspect Fusion**: produce the final query/item representation



Multi-aspect dense retrieval. Kong et al. KDD 2022

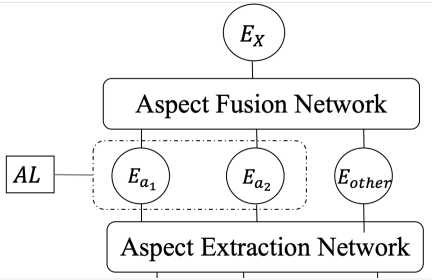
Multi-aspect Dense Retrieval

Three major components:

- **Aspect Extraction**: explicit aspects + a special aspect "other" for implicit semantics

E_X: *Final Representation*

AL: Aspect Learning



- **Experimented on proprietary data**
- **Unreleased source code**

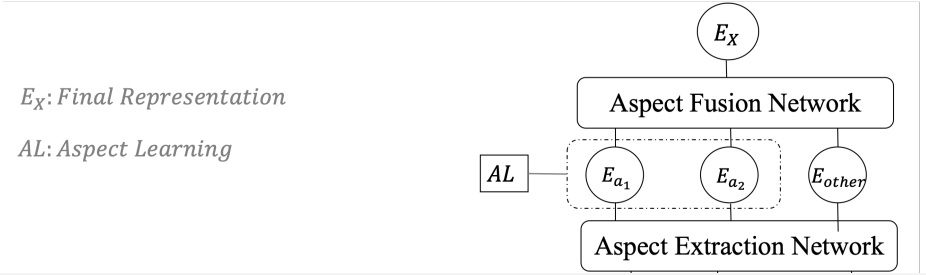
Fail to reproduce on public data!

Multi-aspect dense retrieval. Kong et al. KDD 2022

Multi-aspect Dense Retrieval

Three major components:

- **Aspect Extraction**: explicit aspects + a special aspect “other” for implicit semantics



Why does MADRAL not work?
How to enhance it to work effectively?

Multi-aspect dense retrieval. Kong et al. KDD 2022

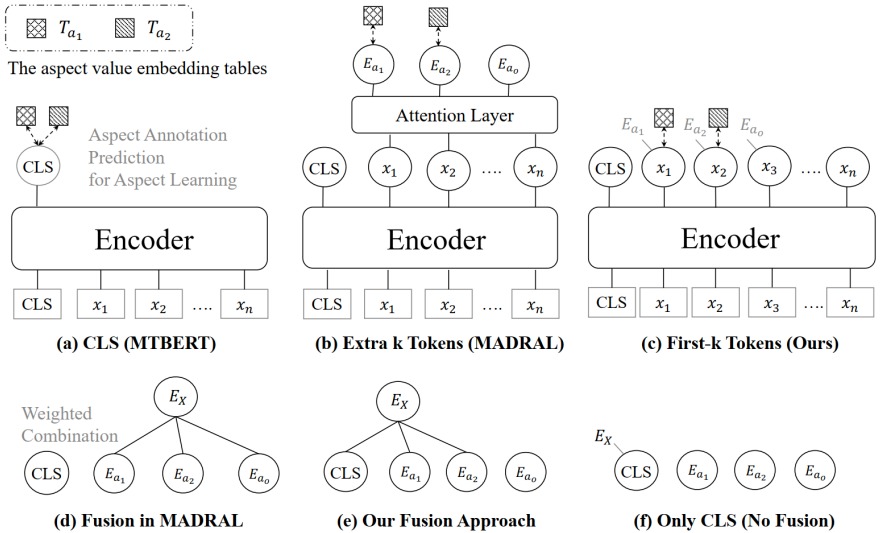
Research Methodology

Are there better choices for the model components?

Multi-aspect Dense Retriever

Three major components:

- **Aspect Extraction**: how to represent aspects
- **Aspect Learning**: predict the value IDs of an aspect
- **Aspect Fusion**: how to fuse the aspect embeddings to produce the final representation



Multi-aspect Dense Retriever

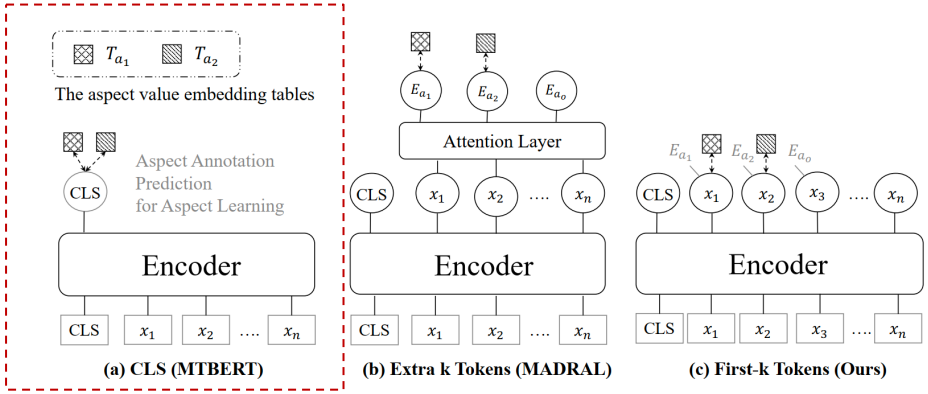
Aspect Extraction

a) **Reusing CLS Token (MTBERT)**

b) Declaring k Extra Embeddings (MADRAL)

c) Reusing First-k Content Tokens

No distinction of importance



Compress information into a single token

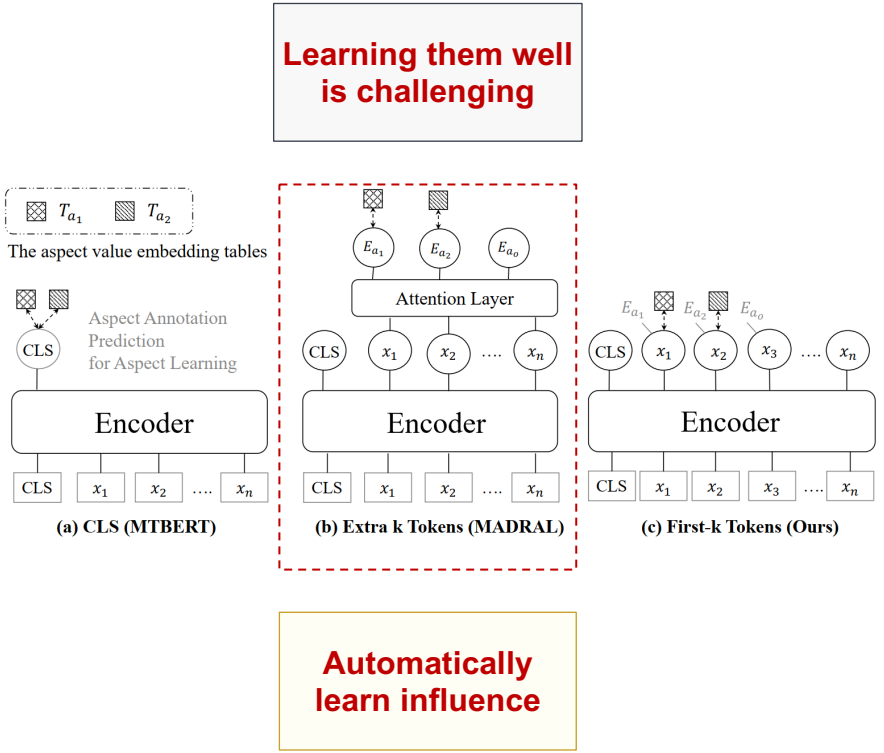
Multi-aspect Dense Retriever

Aspect Extraction

a) Reusing CLS Token (MTBERT)

b) **Declaring k Extra Embeddings (MADRAL)**

c) Reusing First-k Content Tokens



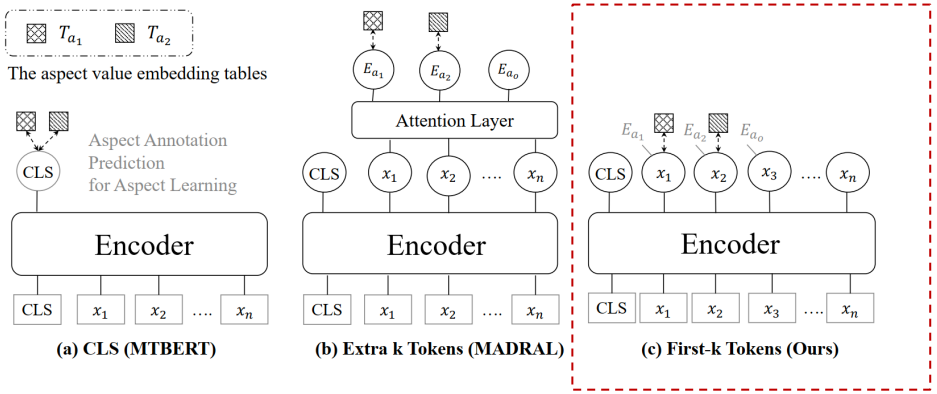
Multi-aspect Dense Retriever

Aspect Extraction

a) Reusing CLS Token (MTBERT)

b) Declaring k Extra Embeddings (MADRAL)

c) **Reusing First-k Content Tokens (Ours)**



Bind to content tokens
+
Decouple the aspects

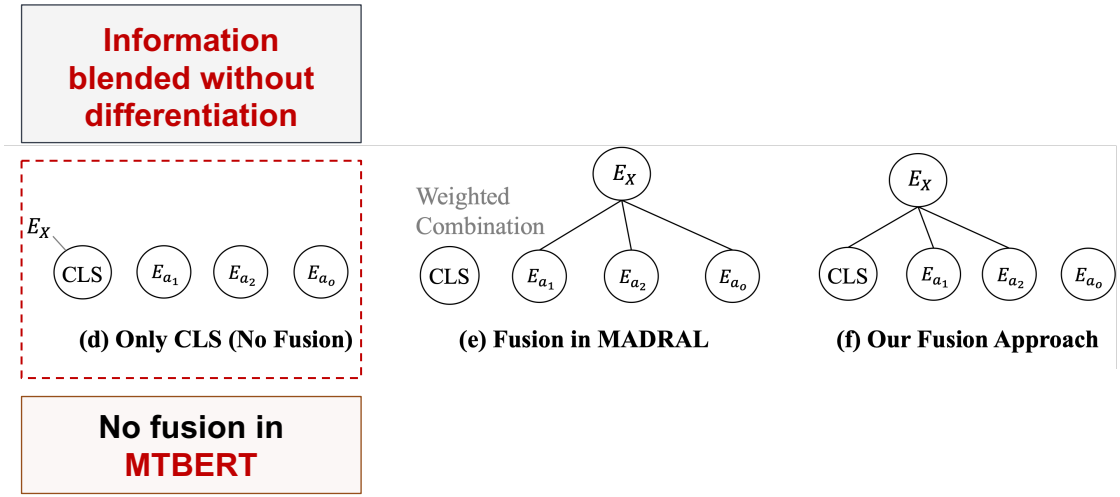
Multi-aspect Dense Retriever

Aspect Fusion (Objects to Fuse)

f) Only CLS (No Aspect Fusion)

d) Aspects and Token “OTHER” (MADRAL)

e) Aspects and Token “CLS” (Ours)



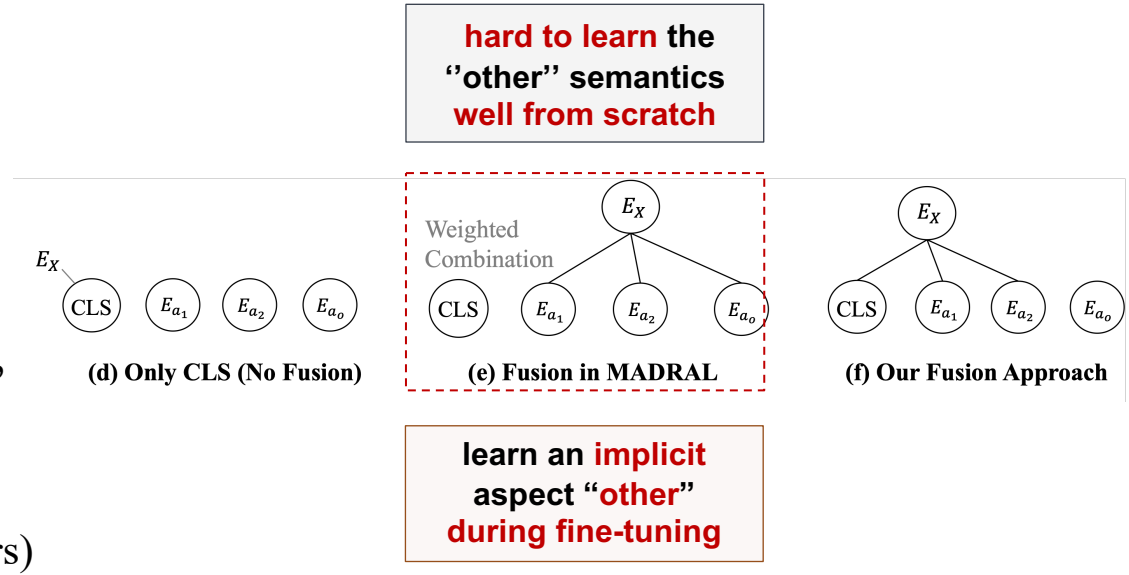
Multi-aspect Dense Retriever

Aspect Fusion (Objects to Fuse)

f) Only CLS (No Aspect Fusion)

d) Aspects and Token "OTHER" (MADRAL)

e) Aspects and Token "CLS" (Ours)



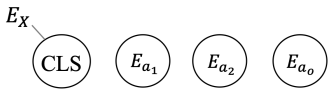
Multi-aspect Dense Retriever

Aspect Fusion (Objects to Fuse)

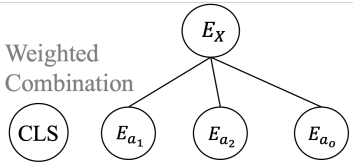
f) Only CLS (No Aspect Fusion)

d) Aspects and Token "OTHER"

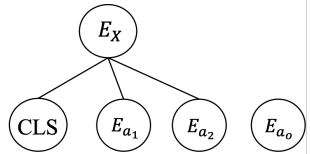
e) Aspects and Token "CLS" (Ours)



(d) Only CLS (No Fusion)



(e) Fusion in MADRAL



(f) Our Fusion Approach

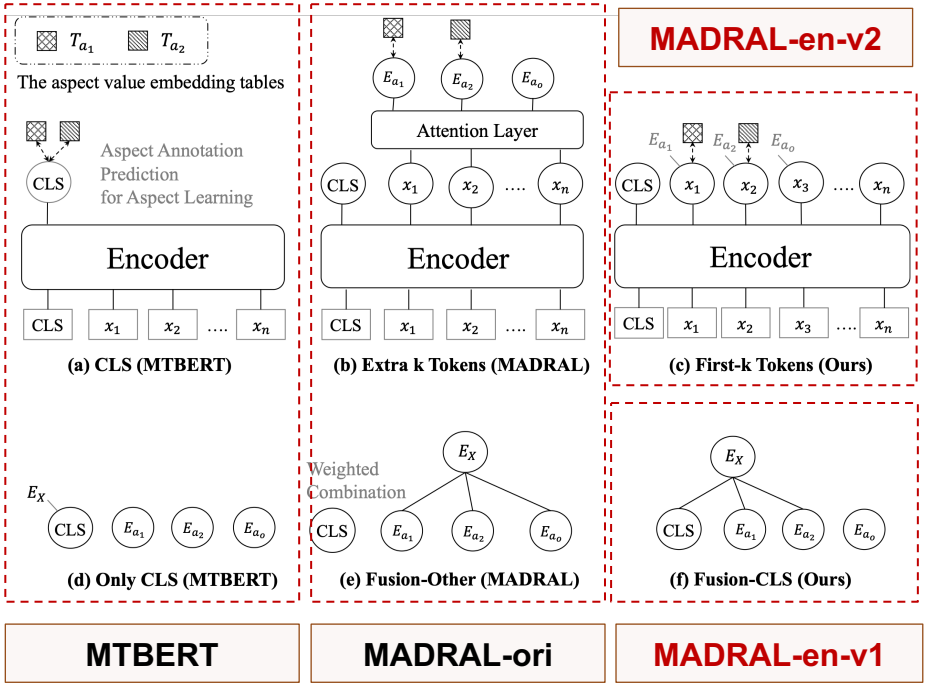
CLS has been sufficiently pre-trained

Reuse CLS to represent explicit content semantics

Multi-aspect Dense Retriever

Three major components:

- **Aspect Extraction**: how to represent aspects
- **Aspect Learning**: predict the value IDs of an aspect
- **Aspect Fusion**: how to fuse the aspect embeddings to produce the final representation



Experimental Results

How do these model variants perform?

Experimental Settings

Multi-Aspect Amazon ESCI Dataset (MA-Amazon)

- 4-level relevance labels: {*Exact, Substitute,Complement, and Irrelevant*}

MA-Amazon				
#train/dev/test-q	#item	item aspects coverage (vocab size)		
		brand	color	category
17k/3.5k/8.9k	482k	94% (5k)	67% (2k)	87% (8k)

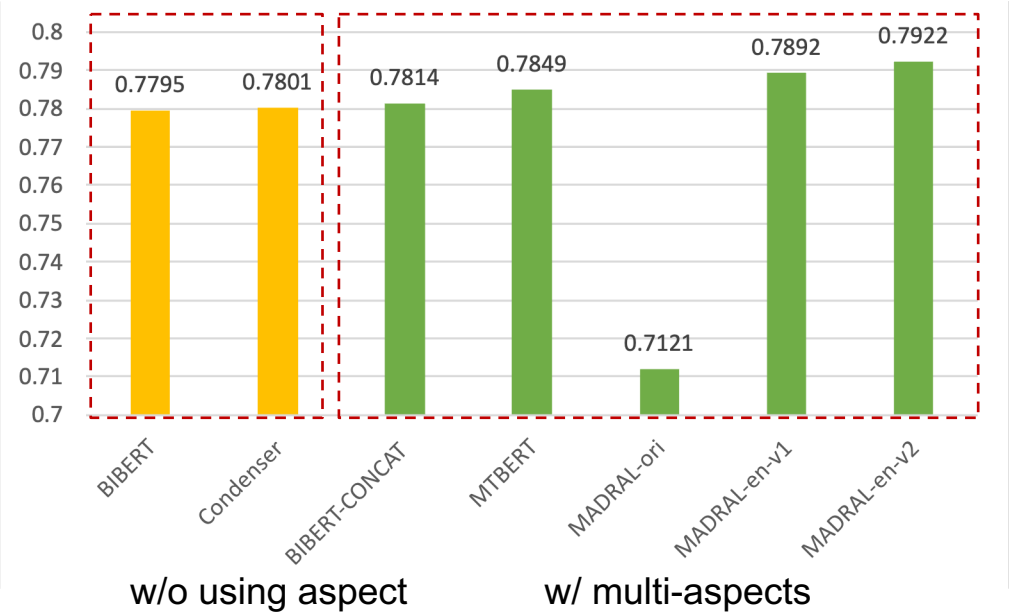
Methods for Comparison

- **BIBERT:** Bi-encoder retriever + Uses "CLS" for encoding
- **Condenser**: An advanced model for textual dense retrieval
- **BIBERT-CONCAT**: Concatenates aspect texts with content as input
- **MTBERT**: Reuses the "CLS" for aspect prediction + No additional fusion
- **MADRAL-ori**: Declares extra aspect embeddings + Fuses aspects with implicit semantics ("OTHER" Emb)

Evaluation Metrics

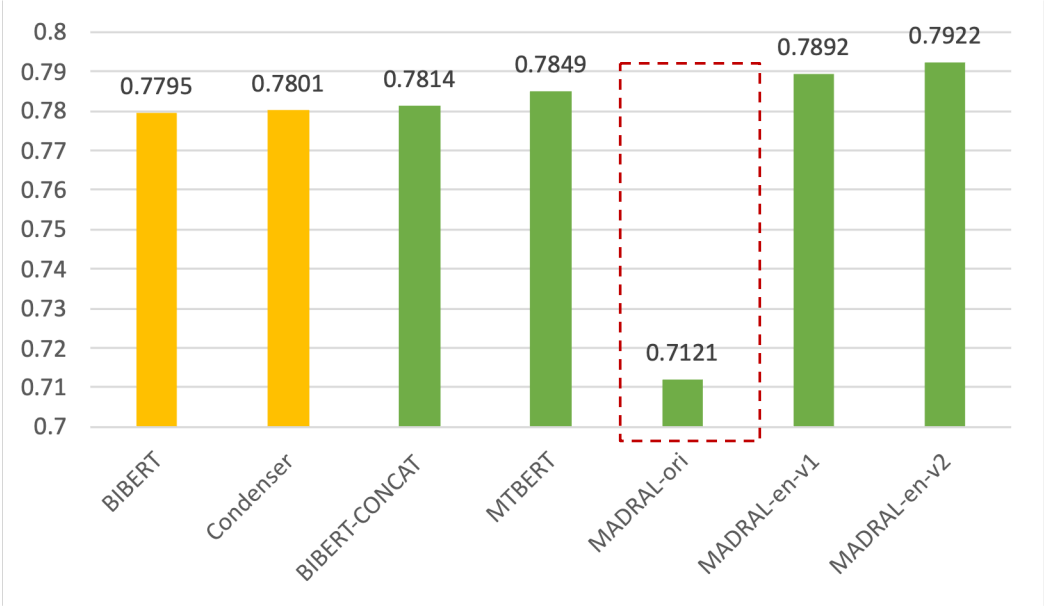
- R@100, **R@500 (for illustration in the slides)**, NDCG@10, and NDCG@50.

(1) Overall Performance



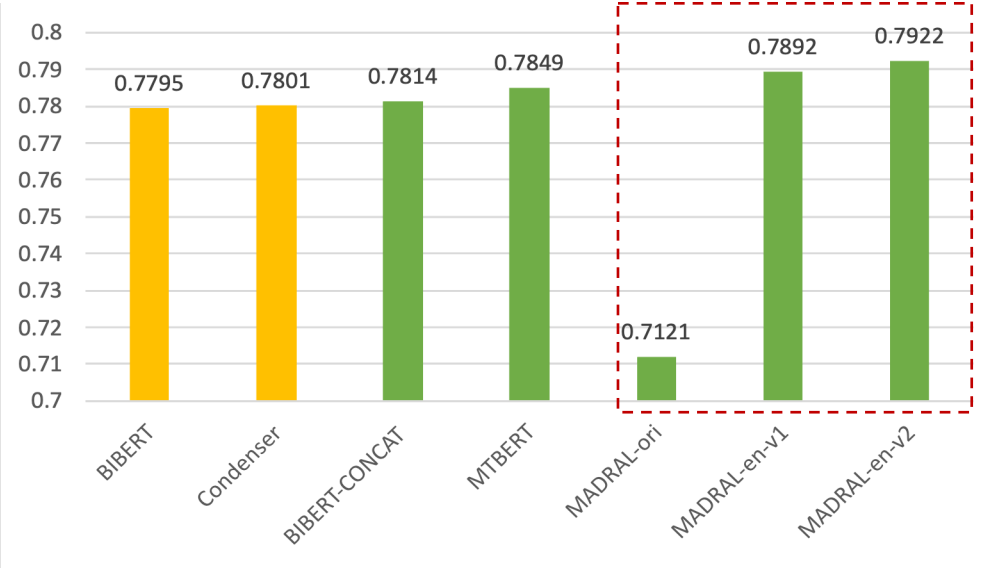
Incorporating aspect information boosts retrieval performance

(1) Overall Performance



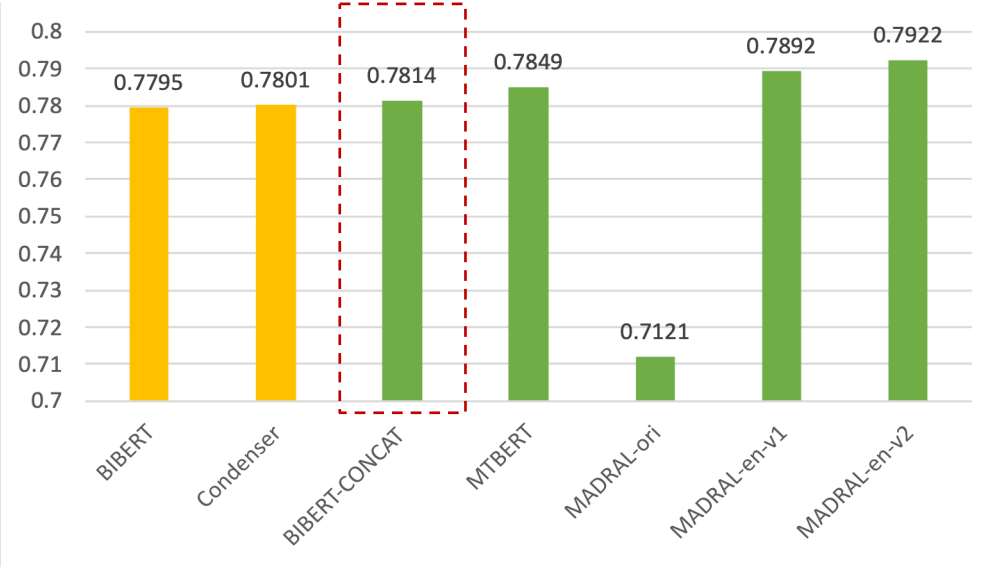
Original MADRAL has significantly worse performance compared to the baselines!

(1) Overall Performance



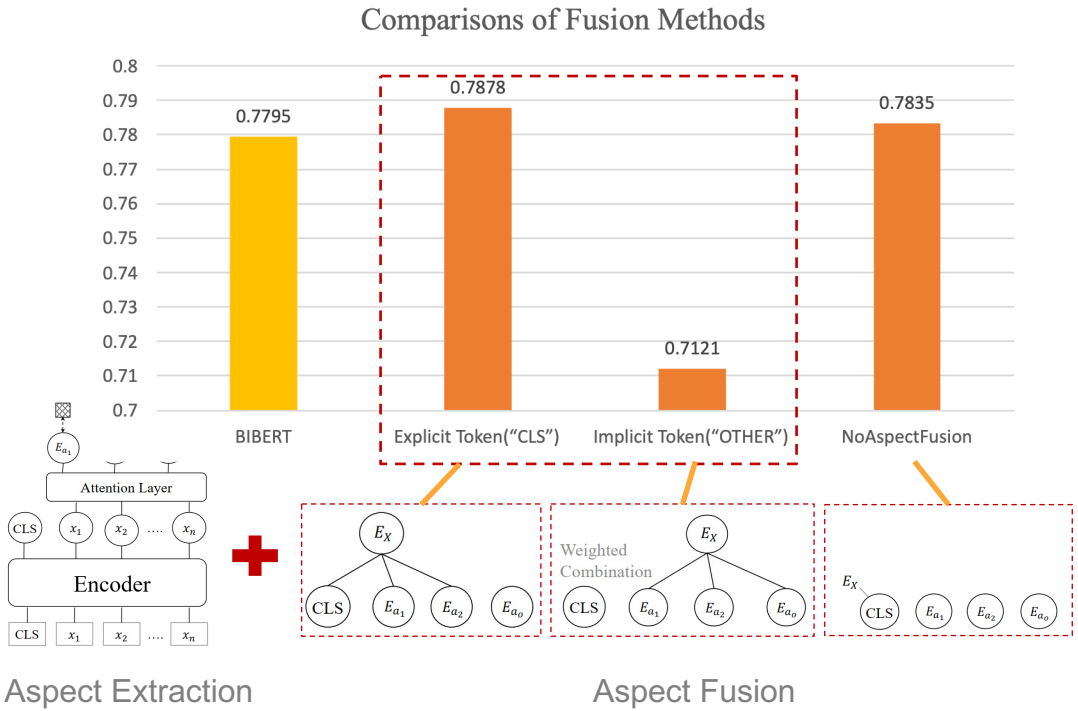
The two MADRAL variants significantly boost the performance of the original MADRAL!

(1) Overall Performance



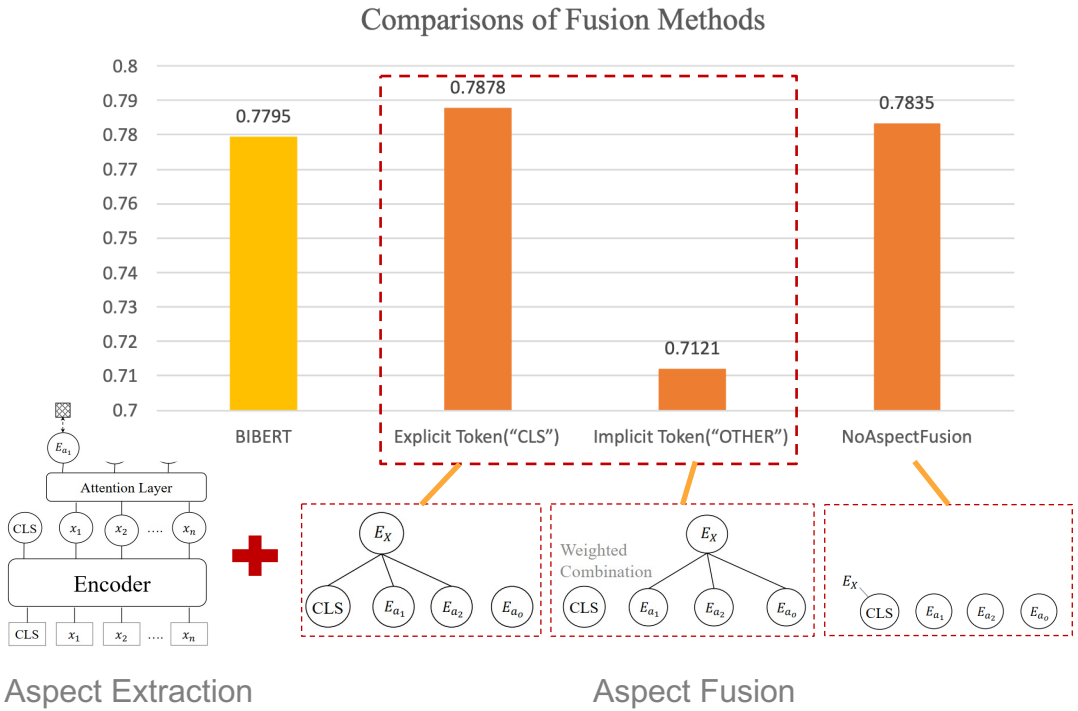
Treating aspects as texts also lead to decent performance (also a promising way of leveraging aspects)

(2) Comparisons of Fusion Methods



Fusion-CLS(Ours) > No Fusion(CLS only) > Fusion-OTHER(Ori)

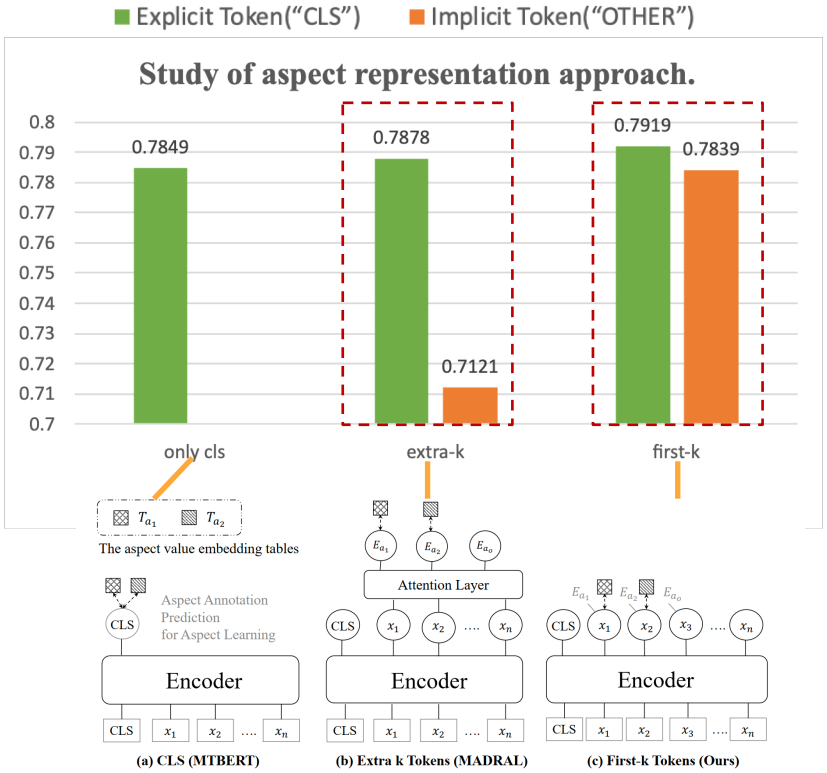
(2) Comparisons of Fusion Methods



Training the "OTHER" token from scratch during fine-tuning is challenging.

Fusion-CLS(Ours) > No Fusion(CLS only) > Fusion-OTHER(Ori)

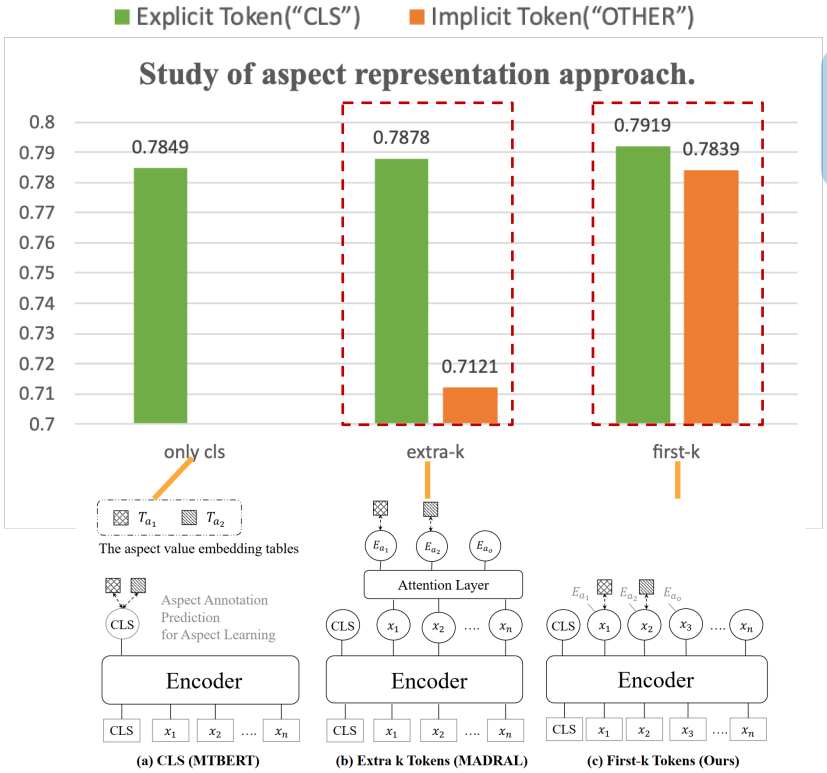
(3) Comparisons of Aspect Representations



First-k (Ours) > extra-k (Ori)

Effectiveness of Reusing Content Tokens: Avoiding the issue of insufficient learning of new parameters.

(3) Comparisons of Aspect Representations



Fusion-OTHER + First-k >> Fusion-OTHER + Extra-k

Attaching aspect learning with existing content tokens reduces the difficulty of learning an implicit "OTHER" embedding during fine-tuning.

(4) Further Analysis

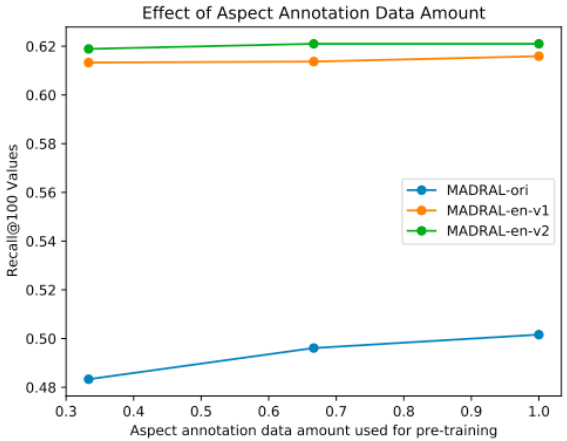


Fig. 4: Effect of Annotation Amount

Aspect Annotation Amount during Pre-training:

- For MADRAL-ori, an increase in annotations improves performance.
- For MADRAL-en-v1/v2, the benefit of more annotations is less pronounced.

Our methods are more annotation data efficient!

Conclusion and Future Work

Conclusion and Future Work

SOTA Model Examination:

- Fail to reproduce the performance of the SOTA model.

Investigation:

- Conducted a comprehensive study of aspect representation and fusion components in multi-aspect dense retrieval.

Key Findings:

- It is easier to learn the aspect embeddings well by reusing existing content tokens than declaring extra aspect embeddings.
- It is better to fuse content semantics (or semantics other than aspects) using “CLS” than a new token "OTHER".

Future Directions:

- Better variants for aspect representation and fusion.
- Treat aspects as text and use it together with aspect value prediction.

Thanks!

Keping Bi

bikeping@ict.ac.cn