Learning List-Level Domain-Invariant Representations for Ranking

Ruicheng Xian[1] Honglei Zhuang[2] Zhen Qin[2] Hamed Zamani[3] Jing Lu[2] Ji Ma[2] Kai Hui[2] Han Zhao[1] Xuanhui Wang[2] Michael Bendersky[2]
[1]University of Illinois Urbana-Champaign [2]Google Research [3]University of Massachusetts Amherst

Overview and Contributions

Revisit domain adaptation for learning to rank via invariant representation learning.

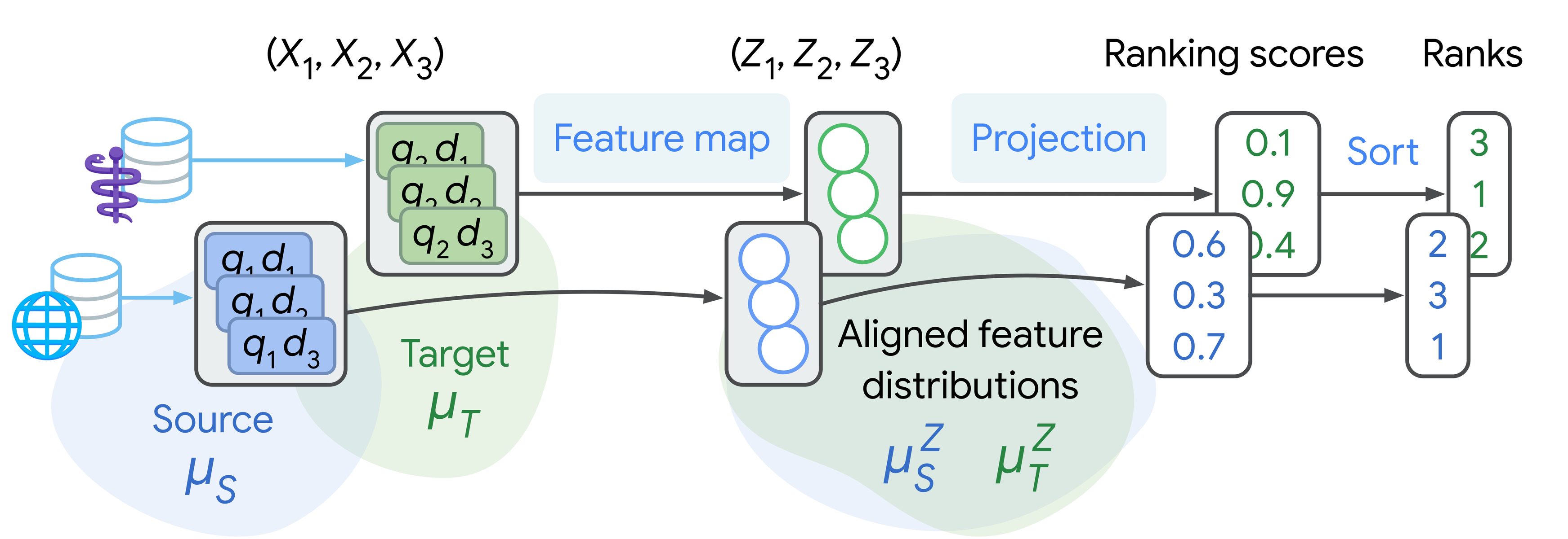
Whereas prior work performs *item-level alignment* [1, 3, 4], we

- propose *list-level alignment*, tailored for ranking;
- establish a domain adaptation generalization bound for ranking based on list-level alignment, and
- demonstrate the its empirical benefits.

Problem and Model Setup, and Invariant Representations

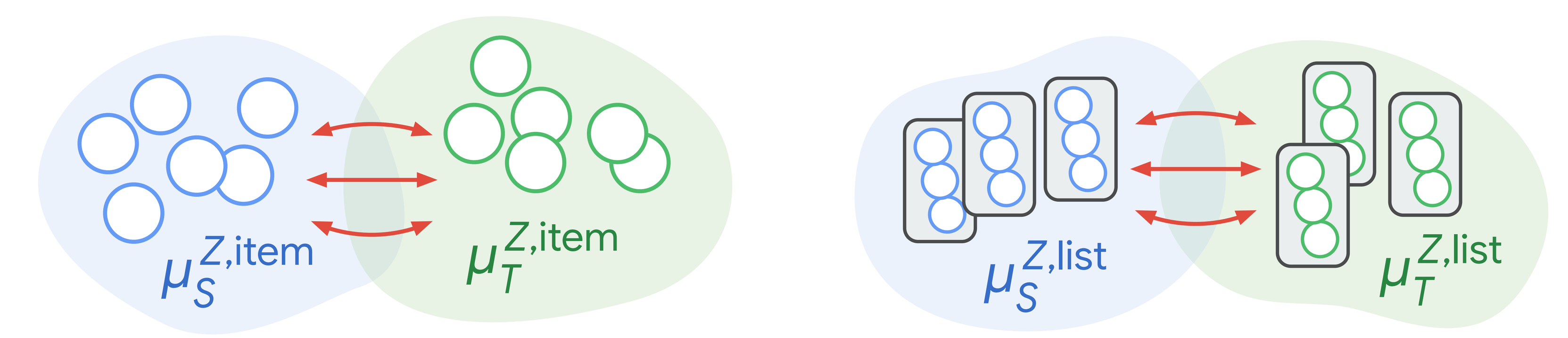
Ranking problems are given by joint distributions μ over **lists** of items $(X_1, \cdots, X_\ell) \in \mathcal{X}$ and relevance scores $(Y_1, \cdots, Y_\ell) \in \mathbb{R}^\ell_{\geq 0}$.

Goal is to obtain a ranking model for a (low-resource, e.g., unlabeled) target domain μ_T, by adapting models trained on a source domain μ_S.



For domain adaptation, we apply *invariant representation learning*, which trains the model to align the source and target domain feature distributions, $\mu_T^Z \approx \mu_S^Z$, where μ^Z is a distribution defined on the vector feature representations, $(Z_1, \cdots, Z_\ell) \in \mathcal{Z} = \mathbb{R}^{\ell \times k}$.

The intuition is that if the source and target data distributions appear similar on the feature space, then models trained on them could transfer across domains.



Item-level alignment List-level alignment

Invariant Representation Learning for Ranking

Item-Level Alignment (ItemDA; prior work)

The implementations in prior work align the distributions of feature vectors (**items**) aggregated from all lists, i.e., $\mu_S^{Z,\text{item}} \approx \mu_T^{Z,\text{item}}$,

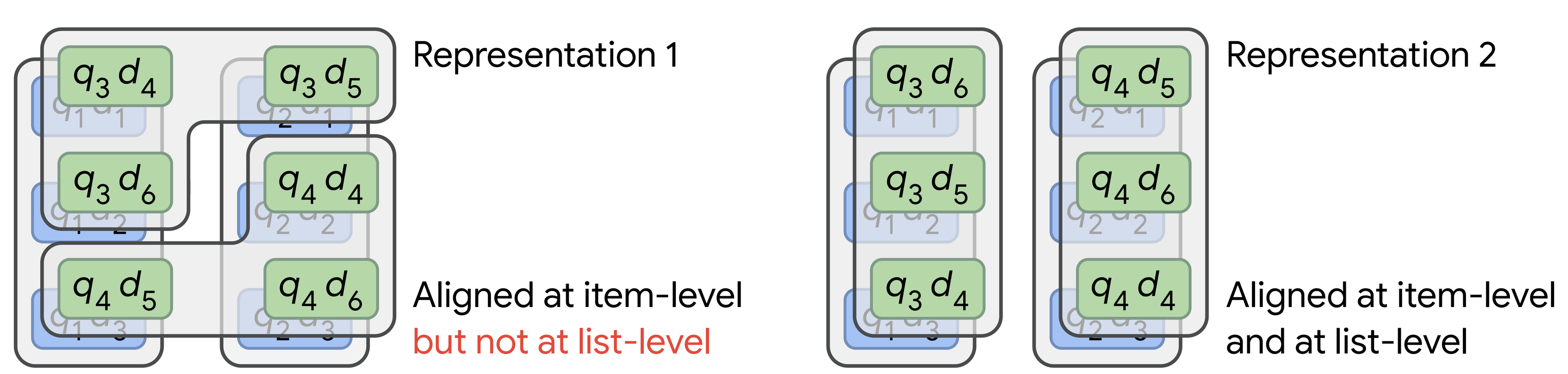
$$\operatorname{supp}(\mu^{Z,\text{item}}) \subseteq \mathbb{R}^k, \qquad \mu^{Z,\text{item}}(v) = \mathbb{P}((Z_1, \cdots, Z_\ell) \ni v),$$

but the list structure on the data is lost from the aggregation step.

List-Level Alignment (ListDA; ours)

To preserve the list structure, we directly align the distributions of **lists** of feature vectors, i.e., $\mu_S^{Z,\text{list}} \approx \mu_T^{Z,\text{list}}$,

$$\operatorname{supp}(\mu^{Z,\text{list}}) \subseteq \mathbb{R}^{\ell \times k}, \qquad \mu^{Z,\text{list}}(z) = \mathbb{P}((Z_1, \cdots, Z_\ell) = z).$$

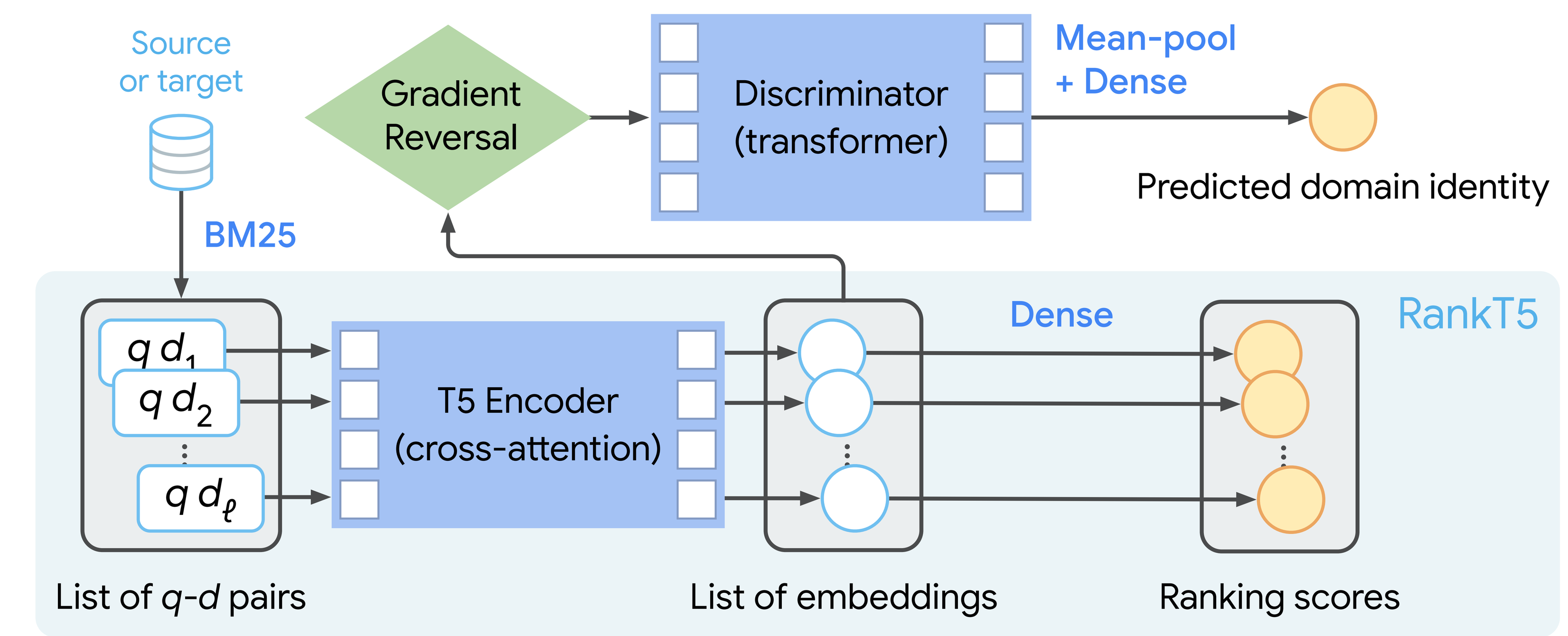


List-level alignment is a stronger requirement than item-level, and is justified by a domain adaptation generalization bound:

Theorem (Instantiated for MRR). Under some Lipschitz assumptions, let $g : \mathcal{X} \to \mathcal{Z}$, then for all scoring models $h : \mathcal{Z} \to \mathbb{R}^\ell$,

$$\mathrm{MRR}_T(h \circ g) \geq \mathrm{MRR}_S(h \circ g) - \Theta(\ell)\, W_1(\mu_S^{Z,\text{list}}, \mu_T^{Z,\text{list}}) - \lambda_g^*,$$

where $\lambda_g^* = \min_{h'}(1 - \mathrm{MRR}_S(h' \circ g) + 1 - \mathrm{MRR}_T(h' \circ g))$ is the minimum joint risk on the learned features (recall $\mathrm{MRR} \in (0, 1]$), and W_1 is Wasserstein distance.



Experiments on Passage Reranking

We adapt RankT5 model [5] from the MS MARCO web search dataset to news and biomedical domains under *unsupervised* setting: on the target domain, only documents are given, and we synthesize queries from documents using a T5 generator [2].

ListDA is compared to zero-shot, ItemDA, and vs. training on pseudolabels generated by the query synthesizer (QGen PL).

Target domain	Method	MAP	MRR@10	NDCG@10
Robust04	BM25	0.2282	0.6801	0.4088
	Zero-shot	0.2759	0.7977	0.5340
	QGen PL	0.2693	0.7644	0.5034
	ItemDA	0.2822*†	0.8037†	0.5396†
	ListDA	**0.2901***†‡	**0.8234***†	**0.5573***†‡
TREC-COVID	BM25	0.2485	0.8396	0.6559
	Zero-shot	0.3083	0.9217	0.8200
	QGen PL	0.3180*‡	0.8907	0.8118
	ItemDA	0.3087	0.9080	0.8142
	ListDA	**0.3187***‡	**0.9335**	**0.8412**†‡
BioASQ	BM25	0.4088	0.5612	0.4653
	Zero-shot	0.5008	0.6465	0.5542
	QGen PL	0.5143*‡	0.6551	0.5643‡
	ItemDA	0.4781	0.6383	0.5343
	ListDA	**0.5191***‡	**0.6666***‡	**0.5714***‡

*Improves upon zero-shot under the two-tailed Student's t-test ($p \leq 0.05$). †Improves upon QGen PL. ‡Improves upon ItemDA.

[1] Cohen et al. Cross Domain Regularization for Neural Ranking Models Using Adversarial Learning. 2018.
[2] Ma et al. Zero-Shot Neural Passage Retrieval via Domain-targeted Synthetic Question Generation. 2021.
[3] Tran et al. Domain Adaptation for Enterprise Email Search. 2019.
[4] Xin et al. Zero-Shot Dense Retrieval with Momentum Adversarial Domain Invariant Representations. 2022.
[5] Zhuang et al. RankT5: Fine-Tuning T5 for Text Ranking with Ranking Losses. 2023.