Transferable and Efficient:
Unifying Dynamic Multi-Domain Product Categorization

Shansan Gong*[1], Zelin Zhou*[1], Shuo Wang[2],
Fengjiao Chen[2], Xiujie Song[1], Xuezhi Cao[2], Yunsen Xian[2], Kenny Q. Zhu[1]

[1]Shanghai Jiao Tong University, [2]Meituan Inc. (*equal contribution)

ACL 2023

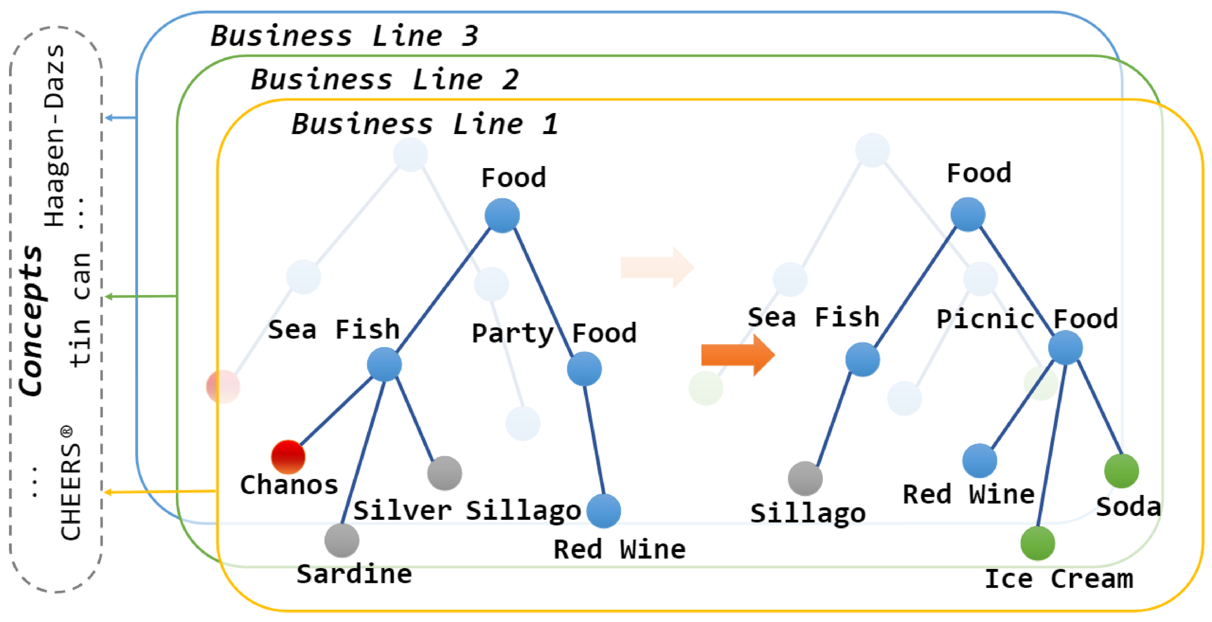
Motivation

We propose Dynamic Multi-Domain Product Categorization (**DMPC**) problem:

- **Multi-domain** taxonomies challenge: e-commerce platforms usually maintain multiple business lines with relatively independent taxonomies;
- **Taxonomy evolving** challenge: with the expansion and reorganization of businesses, each category of taxonomy keeps evolving.

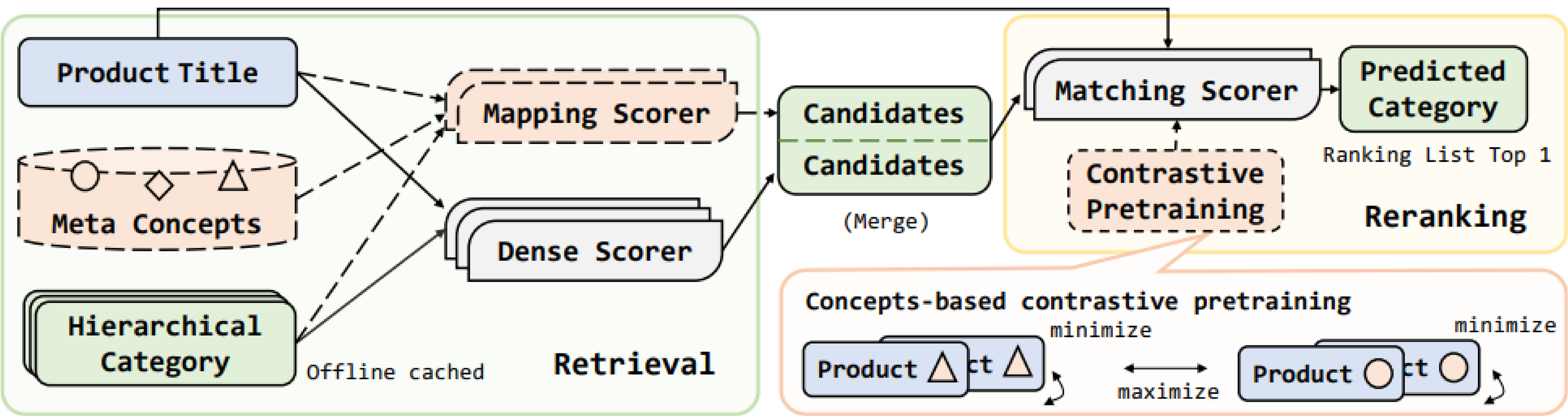
Conventional industry approaches (separately trained classifiers):

- under-utilize the cross-domain data and their shared knowledge;
- raise the expenses of maintenance for different classifiers.



Methods - TaLR framework

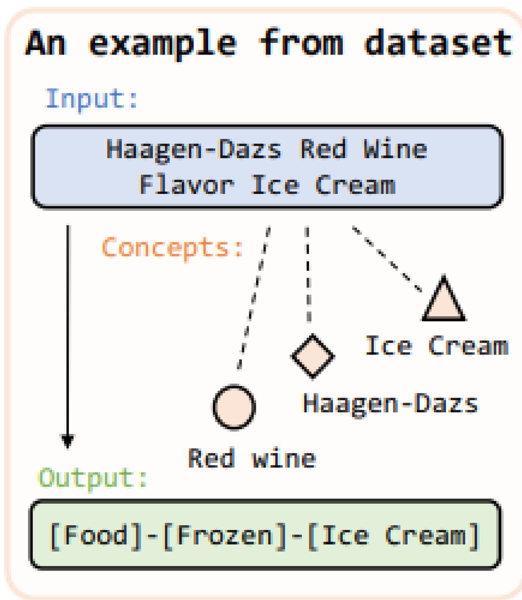
We reformulate the canonical text classification problem as a text relevance matching problem. Our TaLR (Taxonomy-agnostic Label Retrieval) framework is structured into two stages: Retrieval and Reranking. Mapping scorer and contrastive learning are two plug-in modules, both of which are associated with meta concepts.



Datasets

We propose and release Dynamic Multi-Domain Datasets with 3 business lines.

Beyond the category labels, each product title is associated with a list of meta concepts (regarded as cross-domain knowledge)



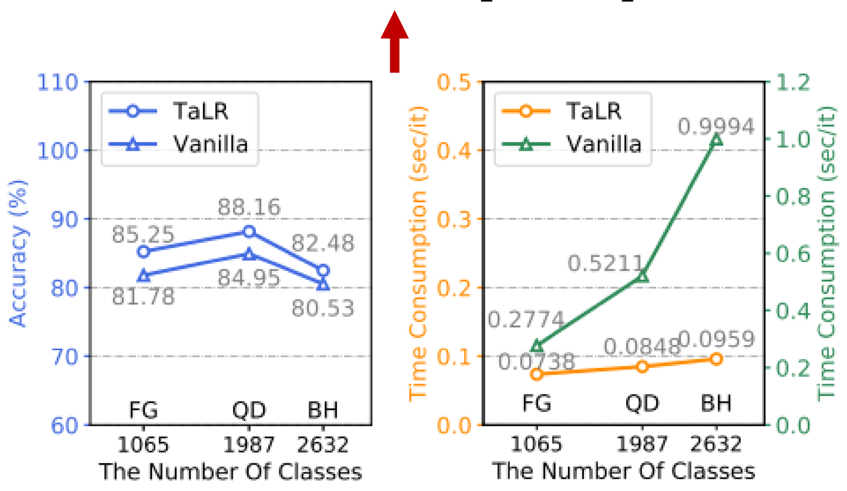
Experiment Results

We compare TaLR and other separately trained baseline models, indicating **knowledge fusion**:

- TaLR achieves even higher accuracy when jointly trained on the mixed multi-domain data
- effectiveness of the two plug-in modules in TaLR framework is orthogonal
- our usage of meta concepts is superior to simple concatenation

We compare inference time for online deployment requirement, indicating **efficiency**:

- inference speed of TaLR is much faster than vanilla model owing to the Retrieval stage
- the time consumption per item of TaLR increases linearly along with the number of classes



The accuracy of TaLR on the new taxonomy

Methods	QD	BH	FG
BERT-matching	9.00	11.23	4.03
BERT-few-shot	43.29	35.19	29.80
TaLR	**60.57**	**65.45**	**62.69**
(-) contrastive	56.71	64.99	60.79
(-) mapping scorer	56.25	64.65	59.29

The accuracy on two dynamic test sets. Δ is the change of accuracy after evolving

Methods	QD−*divide*			QD−*integrate*		
	Before	After	Δ	Before	After	Δ
BERT-matching	6.66	11.95	**+5.29**	13.39	2.23	-11.16
BERT-few-shot	90.51	43.54	-46.96	86.79	50.16	-36.53
TaLR	90.11	**69.71**	-20.40	85.20	**81.48**	**-3.72**

Validate the effectiveness of TaLR to tackle taxonomy evolving challenge:

- TaLR is **robust** to taxonomy evolving
- TaLR can better **transfer** to new taxonomy

The accuracy of baselines and our TaLR framework with variants on static multi-domain datasets

Methods	Overall	QD	BH	FG
Separate models				
TF-IDF&LR	69.51	69.93	68.23	69.95
FastText	74.62	74.01	71.68	80.82
BERT	83.49	84.82	79.93*	84.23
BERT+♠	83.01	86.45	79.02	75.32
HMCN-F-BERT	82.14	83.72	77.09	84.25
HiMatch-BERT	84.08	86.12	77.38	84.19
HiMatch-BERT+♠	83.75	87.26*	77.26	78.53
XR-Linear	76.57	75.27	77.91	78.95
XR-Transformer	84.58*	79.74	79.23	84.58*
XR-Transformer+♠	81.45	85.34	74.59	78.53
(a): TaLR	85.90	87.88	81.92	85.09
Unified model				
BERT Multi-task	68.00	80.27	50.28	44.29
BERT Multi-task+♠	67.79	81.37	49.77	39.83
(b): TaLR	**86.23**	**88.16**	**82.48**	**85.25**
TaLR ablation test				
(c): (b) (-) CL	85.26	86.83	81.75	85.13
(d): (b) (-) MS	84.63	86.59	80.13	84.71
(e): (b) (-) CL&MS	82.82	83.85	79.15	84.71
(f): (b) (-) CL&MS +♠	84.38	87.43	80.64	79.77

♠ concatenate concept text after product title
(-) ablate cretain modules

MS: mapping scorer, CL: contrastive learning

Datasets associated with this paper are released at https://github.com/ze-lin/TaLR.