Centrality-aware Product Retrieval and Ranking

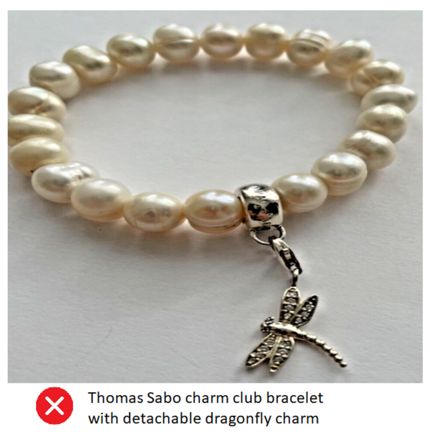
Hadeel Saadany[1], Swapnil Bhosale[2], Samarth Agrawal[3], Diptesh Kanojia[2], Constantin Orăsan[1], and Zhe Wu[3]

[1]Centre for Translation Studies, University of Surrey, UK
[2]Institute for People-Centred AI, University of Surrey, UK
[3]eBay Inc., USA

{hadeel.saadany, s.bhosale, d.kanojia, c.orasan}@surrey.ac.uk, {samagrawal,zwu1}@ebay.com

Introduction

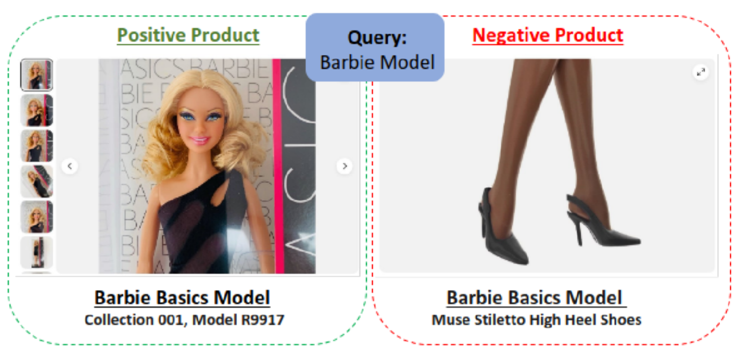
- Enhancing product ranking and retrieval in e-commerce by optimizing existing models.
- User queries carry **ambiguity** and **complexity** ⇒ mismatched intent and retrieval.



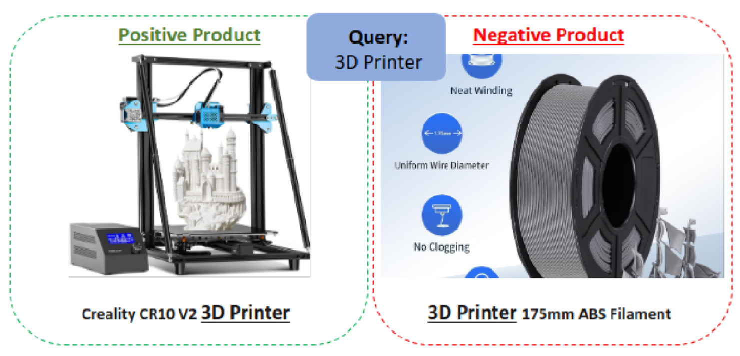
- **Alphanumeric Queries**
  - Queries involving product codes or model numbers, e.g., "`S2716DG`" for a Dell monitor where "S" and "DG" signify specific panel types.
  - Ambiguous product identifiers like "`i5 pc 1tb 16gb 8gb gpu`" that reference specifications but lack complete product details.
- Transformer-based models rely on extensive annotated data but *struggle to accurately capture user intent.*
- Our work proposes a **User-Intent Centrality Optimization (UCO)** to better align models with user intent in product search.

Data & Experiment Setup

- **Dataset-** query-title pairs annotated for **relevance score** and **centrality annotation**.
- Pearson, Kendall and Spearman correlations between the graded relevance score and the binary centrality score are 0.78, 0.73 and 0.77, respectively.
- Relevance scores are assigned on a 1-5 PEGFB scale
  - **1 indicates a Bad match** while **5 indicates Perfect match** (ideal query-title alignment).
- Centrality is annotated as a binary score (1 for central, 0 for non-central) to determine **how closely a title aligns with the query intent**.
- We **curated evaluation splits** from existing data for **four query subsets**.
  - **CQ (Common Queries)**: Consists of general product queries.
  - **CQ-Balanced**: Balanced split with an **equal positive and negative query-title pairs**.
  - **CQ-Common-String**: Contains **queries with exact matches in both relevant and irrelevant titles**, challenging semantic differentiation.



(a) The sub-string "Barbie Model" is a part of both positive and negative product titles.



(b) The sub-string "3D Printer" is a part of both positive and negative product titles.

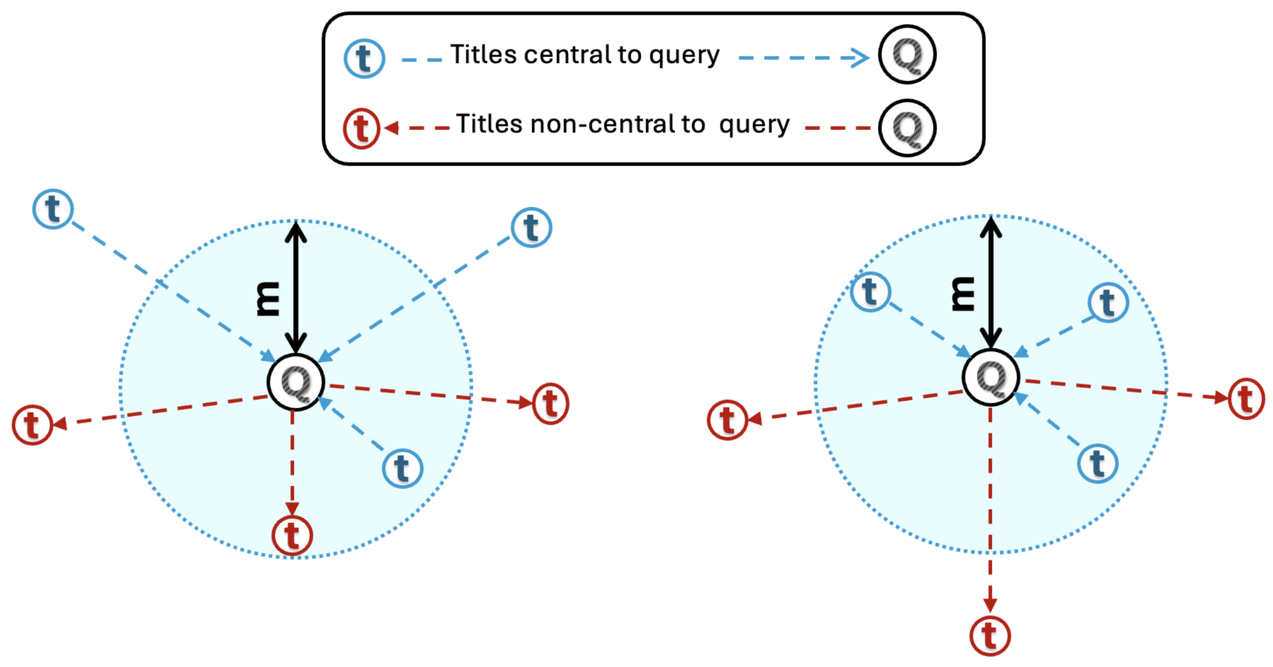
  - **CQ-Alphanumeric**: Focuses on alphanumeric queries, **such as product codes or model numbers**, where minor changes can significantly impact retrieval accuracy.

Evaluation Split	Corpus Size	Dev Queries	Test Queries
CQ	$187,469$	$5,776$	$17,325$
CQ-Balanced	$46,561$	$5,776$	$17,325$
CQ-Common-String	$12,508$	$2,117$	$6,351$
CQ-Alphanumeric	$162,115$	$4,111$	$12,333$

- **eBERT** and **eBERT-siamese** models are used as encoder backbones, pre-trained on eBay's item data combined with general domain text.
- Fine tuned for a maximum of 10 epochs, with a Batch size of 32 using AdamW (learning rate $\rightarrow 2e-05$, weight decay $\rightarrow 0.01$, with **cosine similarity** as evaluation function.
- Ranking evaluation metrics include **Precision@k**, **Recall@k**, **NDCG@k**, and **Mean Reciprocal Rank (MRR)** to measure the quality and ranking accuracy.

User-Intent Centrality Optimization (UCO) Approach

- The proposed User-Intent Centrality Optimization (UCO) enhances product title ranking by aligning retrieval with buyer intent.
- UCO uses a **Dual Loss-based Optimization** to manage hard negatives in query-title pairs (as shown in figure):
  - **Multiple Negative Ranking Loss (MNRL)**: Increases the distance between positive and negative pairs in the model's embedding space.
  - **Online Contrastive Loss (OCL)**: Optimizes for hard positives (distant in embedding space) and hard negatives (closer to positive pairs).



- The ablation study demonstrates that **combining both loss functions, MNRL and OCL, enhances the model's performance**, as the dual-loss approach improves ranking quality.
  - Employed individually, MNRL seems to outperform OCL in both metrics.

Loss Function	NDCG@5	MRR@10
MNRL	0.7139	0.7899
OCL	0.5497	0.6559
MNRL + OCL (Dual-Loss)	**0.7488**	**0.8189**

- **Fine-tuning** process:
  - UCO fine-tunes the model on **binary classification task for centrality**.
  - Dual loss ensures a **balance between relevance** (matching the query intent) and **centrality** (most typical product titles).

Results

- UCO helps distinguish between titles semantically relevant but non-central and central to user intent, mitigating challenges posed by ambiguous/alphanumeric queries.

Encoder	UCO	Precision@k (↑) 3	5	10	Recall@k (↑) 3	5	10	NDCG@k (↑) 3	5	10	MRR (↑) @10
***CQ* test**											
BERT	No	16.20	13.03	8.93	11.31	14.41	18.83	0.1912	0.1818	0.1833	0.2771
eBERT	No	20.71	17.25	12.54	14.46	19.19	26.26	0.2392	0.2330	0.2430	0.3415
	Yes	**64.76**	**55.74**	**39.22**	**49.63**	**63.92**	**79.65**	**0.7439**	**0.7488**	**0.7672**	**0.8189**
eBERT (siam)	No	55.25	48.33	34.90	42.36	56.09	72.22	0.6315	0.6428	0.6704	0.7263
	Yes	**66.25**	**57.16**	**40.20**	**51.18**	**65.79**	**81.66**	**0.7635**	**0.7698**	**0.7886**	**0.8347**
***CQ-balanced* test**											
BERT	No	7.13	4.94	2.95	21.26	24.58	29.33	0.1824	0.1961	0.2115	0.1862
eBERT	No	9.72	6.94	4.22	29.02	34.58	42.07	0.2428	0.2657	0.2899	0.2495
	Yes	**28.57**	**18.15**	**9.50**	**85.40**	**90.42**	**94.62**	**0.7851**	**0.8059**	**0.8197**	**0.7789**
eBERT (siam)	No	25.99	16.68	8.89	77.66	83.08	88.59	0.6888	0.7112	0.7291	0.6784
	Yes	**29.19**	**18.39**	**9.58**	**87.26**	**91.58**	**95.43**	**0.8046**	**0.8225**	**0.8351**	**0.7965**
***CQ-common-str* test**											
BERT	No	9.41	6.31	3.65	28.15	31.47	36.35	0.2532	0.2669	0.2828	0.2579
eBERT	No	12.62	8.64	5.00	37.79	43.10	49.92	0.3272	0.3491	0.3714	0.3315
	Yes	**32.03**	**19.58**	**9.92**	**95.84**	**97.65**	**98.87**	**0.9091**	**0.9166**	**0.9206**	**0.8979**
eBERT (siam)	No	29.93	18.76	9.68	89.57	93.58	96.50	0.8194	0.8361	0.8456	0.8063
	Yes	**32.12**	**19.64**	**9.92**	**96.11**	**97.94**	**98.93**	**0.9117**	**0.9193**	**0.9226**	**0.9003**
***CQ-alphanum* test**											
BERT	No	20.54	16.65	11.47	13.45	17.32	22.82	0.2333	0.2176	0.2226	0.3350
eBERT	No	23.35	19.54	13.77	15.53	20.76	27.85	0.2630	0.2516	0.2617	0.3739
	Yes	**64.58**	**57.27**	**40.35**	**44.05**	**59.97**	**77.00**	**0.7119**	**0.7094**	**0.7344**	**0.8018**
eBERT (siam)	No	60.67	54.10	38.54	41.32	57.10	74.20	0.6652	0.6654	0.6951	0.7618
	Yes	**67.10**	**59.70**	**41.81**	**46.07**	**62.72**	**79.76**	**0.7375**	**0.7371**	**0.7609**	**0.8171**

Conclusion

- Our proposed UCO method effectively **improves product search relevance** by **aligning model rankings with user intent**, showing **consistent performance gains across evaluation metrics**.
- UCO's dual-loss approach **optimizes the embedding space** to **better handle** challenges, such as **ambiguous and alphanumeric queries**, ensuring that results align more closely with user expectations.
- Future work will explore **explainable product retrieval for complex queries** and **leveraging GenAI to expand challenging query structures** and align them with user intent.