Improving Text-based Similar Product Recommendation for Dynamic Product Advertising at Yahoo

Xiao Bai, Lei Duan, Richard Tang, Gaurav Batra and Ritesh Agrawal

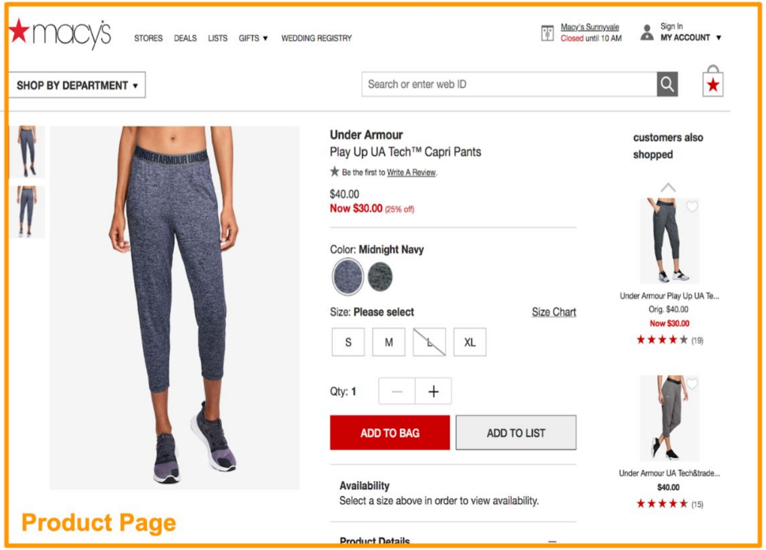
Yahoo Research & Yahoo

CIKM'22, October 17-21, 2022, Atlanta, Georgia, USA

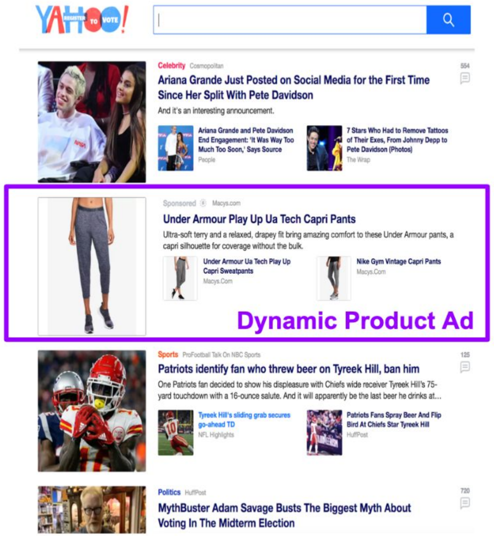
Dynamic Product Advertising

Personalized ad experience for e-commerce users

Advertiser Website



Publisher Website



Similar Product Recommendation

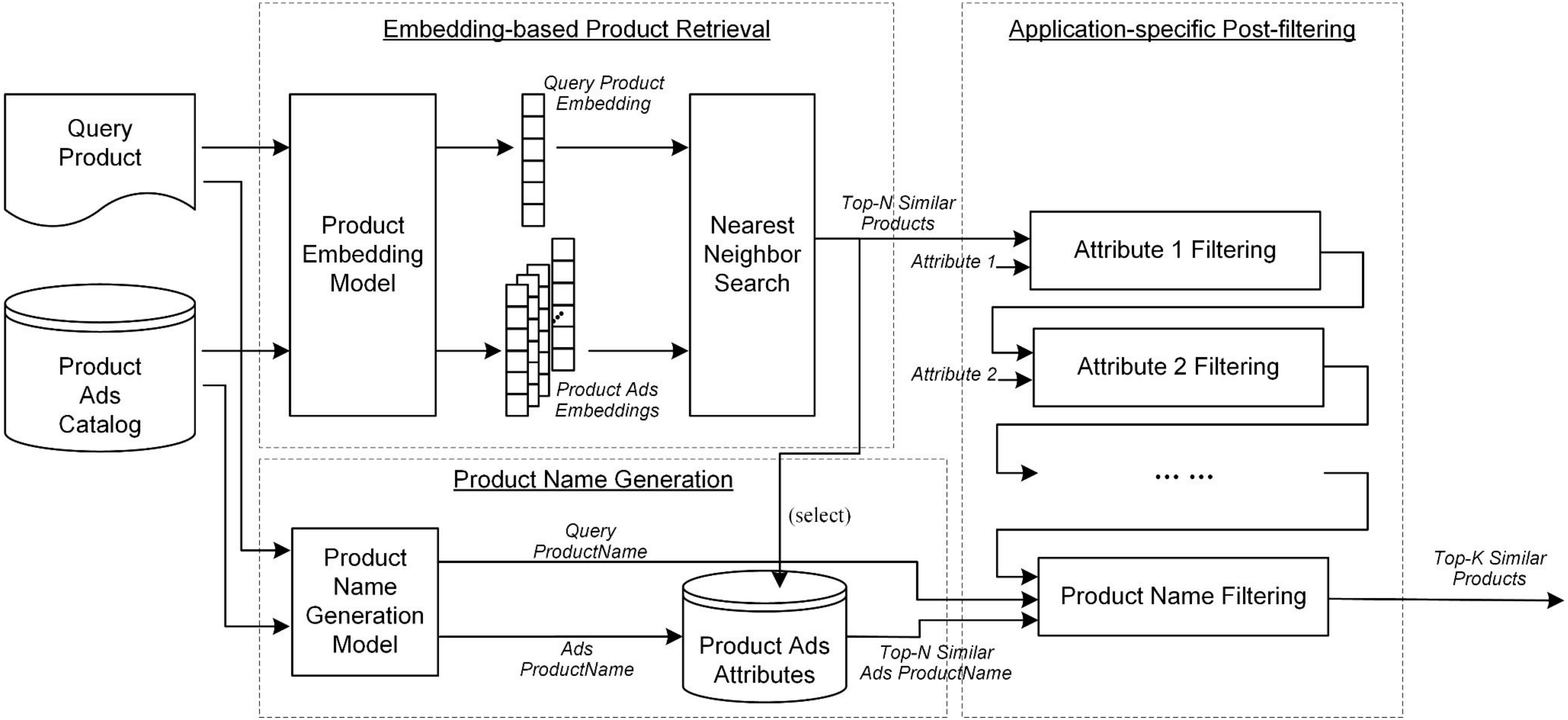
Crucial to the success of Dynamic Product Advertising

- Similar products as alternatives in an ad carousel
- Similar products as substitutes of out of stock products
- Similar products to bring new users to advertiser websites

Challenges and Contributions

Key Challenges	Our Solutions
Application specific models are costly →	**Retrieve and filter paradigm** - Embedding-based semantic retrieval - Application-specific filtering
Clicks are very weak similarity signals →	Human annotated data for semantic retrieval - **Editorial guidelines** for general purpose product similarity
Embedding-based semantic retrieval may fail to capture key product aspects →	Product type as a post retrieval filter - **Product name generation** from text

System Overview



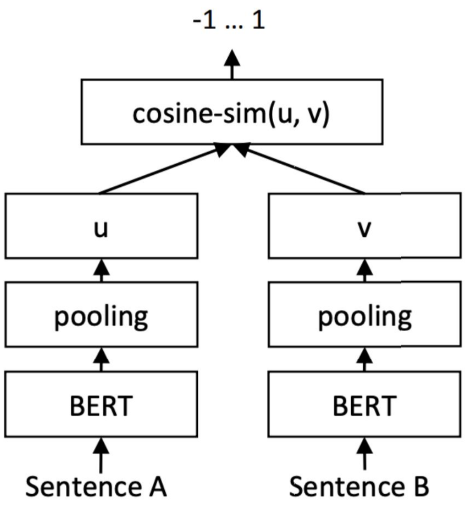
yahoo!

yahoo! research

Embedding-based Product Retrieval

Transformer-based Siamese Network (aka Sentence-BERT [1])

- Fine-tune with editorial labeled dataset
  - Excellent/Good/Fair as positive (+1)
    - Excellent: exactly the same products
    - Good: Same type but differ in brand, model, etc.
    - Fair: Same category but differ in major characteristics (bike vs. electric bike)
  - Bad as negative (-1)
    - Different types of products (boots vs. sandals)
    - Gender/age mismatch
- Product titles as inputs
- Mean-pooling of hidden representations as product embedding
- Mean-squared loss on cosine similarity



Reprinted from [1]

[1] Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In EMNLP'19.

Product Name Generation

Problem definition

- Given a product, generate all the valid product names, from its *title* and *GPT* (Google Product Taxonomy) category
- Product name
  - A small number of words that describe the type of a product.
  - Core product + at most one modifier that is not an existing product attribute (e.g., age, gender, brand, etc.)

*Title: Star K Heart Shape 8mm Created Sapphire Antique Vintage Style Solitaire **Engagement** Promise **Ring***

GPT: Apparel & Accessories > Jewelry > Rings

Product names:
- ring
- engagement ring

modifier
core product

Background

Product attribute identification in the literature

- Multi-class classification for each attribute
    - Works well for attributes with a small number of fixed values, e.g., gender, age
    - Cannot predict new attribute values, e.g., product name, brand
- Sequence tagging for named entity recognition
    - Typical models: CRF, BiLSTM, BiLSTM-CRF, etc.

Title: Star K Heart Shape 8mm Created Sapphire Antique Vintage Style Solitaire ***Engagement*** *Promise* ***Ring***

Tags: O O O O O O O O O O O ***B I I***

- Cannot extract product names that are not a continuous text span

Product name identification as a keyphrase generation task in this work

Pre-trained Transformer for Product Name Generation

Fine-tuning pre-trained Transformers with seq2seq objectives

- Input format

[*SOS*] title [*SEP*] GPT **[*EOS*]** $pn_1;pn_2...;pn_t$ **[*EOS*]**

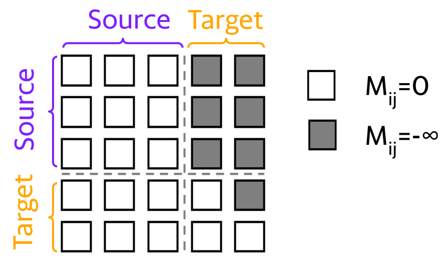
Source: title and GPT

Target: product names as a concatenated list w.r.t. One2Seq schema

- Self-attention mask[1] for product name generation

$$A = softmax\left(\frac{QK^T}{\sqrt{d_k}} + M\right) V,$$

Self-attention mask



- Fine-tuning objective
    - Maximize the likelihood of the target sequence given the source sequence

[1] Hangbo Bao et al. UniLMv2: Pseudo-Masked Language Models for Unified Language Model Pre-Trainings. In ICML'20.

Evaluation of Embedding-based Product Retrieval

Datasets

- Train and Dev: 500K product-product pairs
- Test: 10K product-product pairs for head, torso, tail products respectively
- 5 random similar products per product, labeled by professional editors as Excellent/Good/Fair/Bad

Metrics: NDCG@5

CNN for low dimensional representation

Model	Embedding Size	Head	Torso	Tail
Production (GBDT)	NA	0.8782	0.7713	0.7366
CLSM [35]	128	0.8756	0.7754	0.7388
EPR (Our model)	128	0.8837	0.7843	0.7746
EPR (Our model)	768	0.8836	0.7845	0.7750

Random projection

Evaluation of Product Name Generation

Datasets

- Train and Dev: 55K Random + 2.5K Difficult + 2.5K Ambiguous products
- Test: 3K Random; 5K Easy; 2.5K Difficult; 2.5K Ambiguous products
- Product difficulty determined by the number of product names given by production model

Metrics

- F1@M - M is the actual number of product names being generated/extracted

Model	Random	Easy	Difficult	Ambiguous
Production-CRF [24]	0.381	0.503	0.000	0.357
BiLSTM-CRF [18]	0.602	0.616	0.375	0.558
SEG-NET [3]	0.436	0.432	0.392	0.416
PNG (Our model)	0.683	0.628	0.583	0.659

Transformer trained from scratch

Extractive

Extractive & Generative

Generative

Product Name as a Post Filter

Filter out if:

- No common product name between two products

Datasets

- 1000 random product with top-5 similar products

Approach	Coverage	Excellent	Good	Fair	Bad
EPR	87.8%	7.4%	85.0%	2.4%	5.3%
EPR+Production	53.4%	8.8%	86.9%	1.3%	3.1%
EPR+PNG	81.0%	8.1%	86.0%	2.0%	3.9%

Application 1: Substitute of Out-of-Stock Products

Setting

- Retargeting: showing users the same products they were interested in
- Top-1 similar product as substitute for each out-of-stock Retargeting product
- Application specific filter:
    - Different advertisers

Online A/B testing results

Impression	CPM	CTR
+3.8%	+2.4%	+0.69%

Application 2: Prospecting via Similar Products

Setting

- Prospecting: showing users products they might be interested in based on implicit signals
- Top-3 similar products for each product eligible for Retargeting
- Application specific post filter:
  - Same titles if different advertisers

Online A/B testing results

Impression	CPM	CTR
+2.2%	+1.1%	+0.0%

Conclusions

Text-based similar product recommendation

- Transformer-based Siamese product retrieval model improves tree-based and CNN-based models
- Product name as post filtering improves embedding-based product retrieval
- Transformer-based product name generation model improves extractive models

Applications in Yahoo Dynamic Product Advertising

- Retrieve and filter paradigm easily supports similar applications
- High quality similar products improve ad impression and revenue in two applications

Thanks!

Questions / Suggestions?

Improving Text-based Similar Product Recommendation
for Dynamic Product Advertising at Yahoo

xbai@yahooinc.com