

Product Collection Recommendation in Online Retail



Motivation

- Behavioral data shows that users buy more than one products per transaction
- Bundle recommendation**, i.e., recommending a set of products that can all be purchased together, is of benefit to sellers and buyers
- Oftentimes, the user is looking to to buy a set of products that have **common** theme, e.g., bathroom remodeling



Problem Definition & High-level Approach

Collection recommendation problem: products belong to the same collection if they satisfy a set of constraints such as:

Match Styles

Constitute solutions for the same project

Products are not interchangeable

Our Approach Combines:

Content information, e.g. hierarchies, title, description

Title: Banbury Centerset
2-Handle High-Arc Bathroom
Faucet in Brushed Nickel

Description: A timeless look to your bathroom with the Banbury 2-Handle High-Arc Bathroom Faucet from MOEN ...

Hierarchy: Bath > Bathroom Faucets > Bathroom Sink Faucets > Centerset Bathroom Sink Faucets

Transactional data, e.g. co-purchases



Domain Knowledge, e.g. collection relationships for a small number of products



Proposed Solution

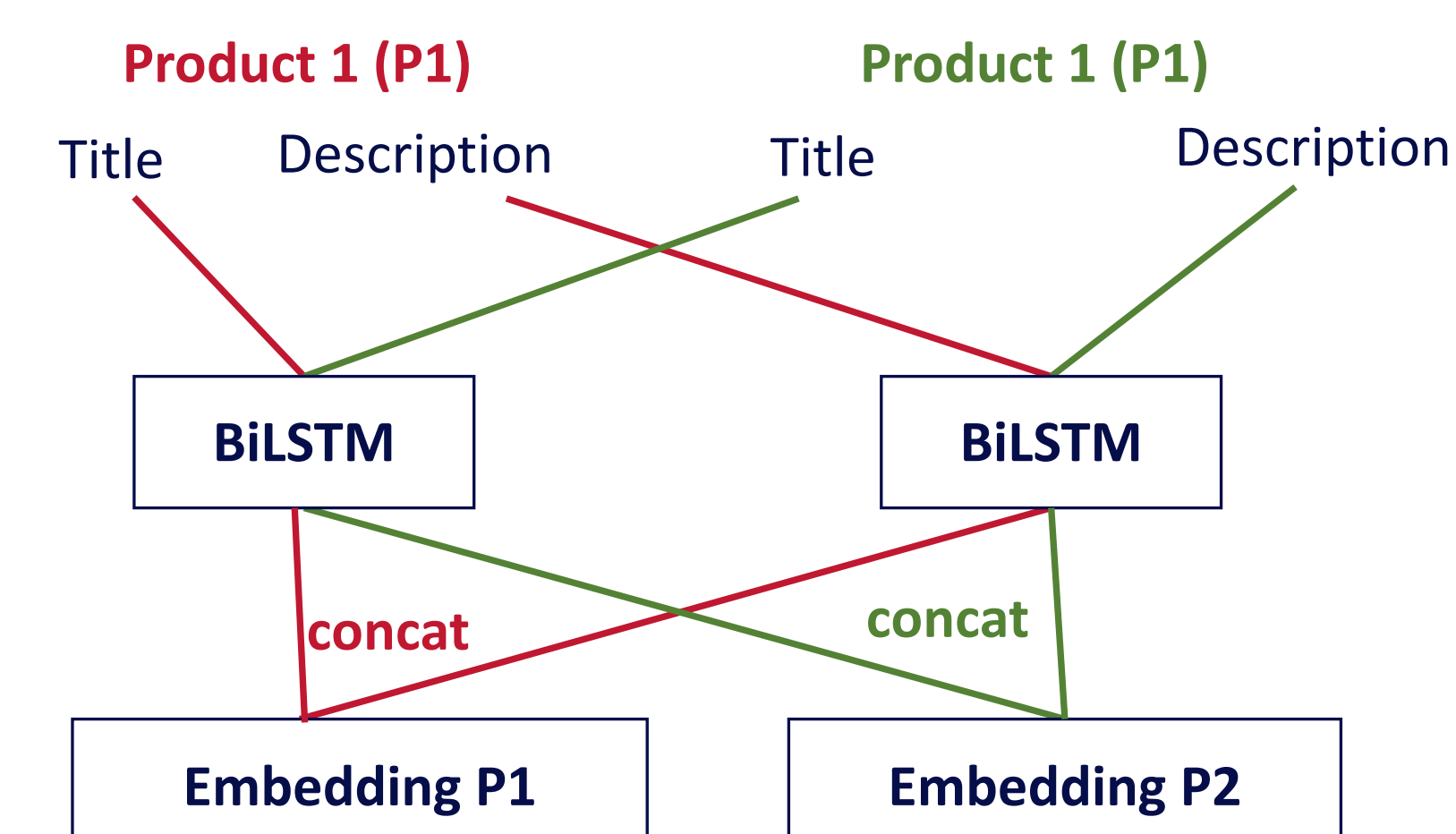
Step One

Generate candidate sets of products that potentially belong to the same collection



Step Two

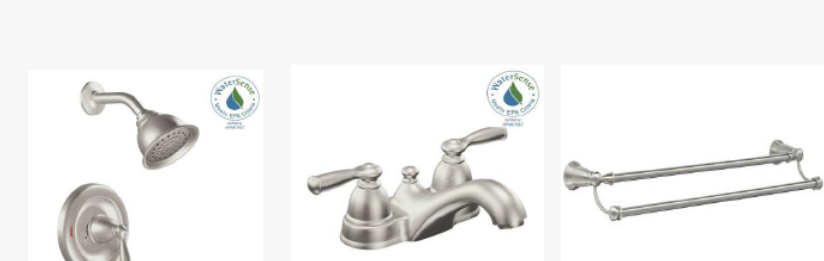
Identify products from the candidate set that indeed belong to the same collection



Finally, we use **cosine similarity** to compute similar embeddings between products. Similar embeddings correspond to products that belong to the same collection.

Examples

Shower Faucets
Bathroom sink faucets
Towel Bars



Patio rocking chair
Outdoor ottomans
Outdoor side tables



Evaluation Set Up

01 Competing methods

- Collections from domain experts (Experts)
- Proposed Method 1 (Method1)
- Proposed Method 2 (Method2)

02 Datasets

- Bathroom: 487 anchor products
- Patio: 369 anchor products

03 Evaluation process

- For each anchor product, for each method, we generated 5 collection recommendations
- Used Home Depot's validation team to validate whether each recommendation was relevant

Evaluation Results

	Bath	Patio
Method	Accuracy (SD)	Accuracy (SD)
Experts	0.51 (0.32)	0.76 (0.23)
Method 1	0.75 (0.15)	0.83 (0.24)
Method 2	0.82 (0.13)	0.83 (0.24)

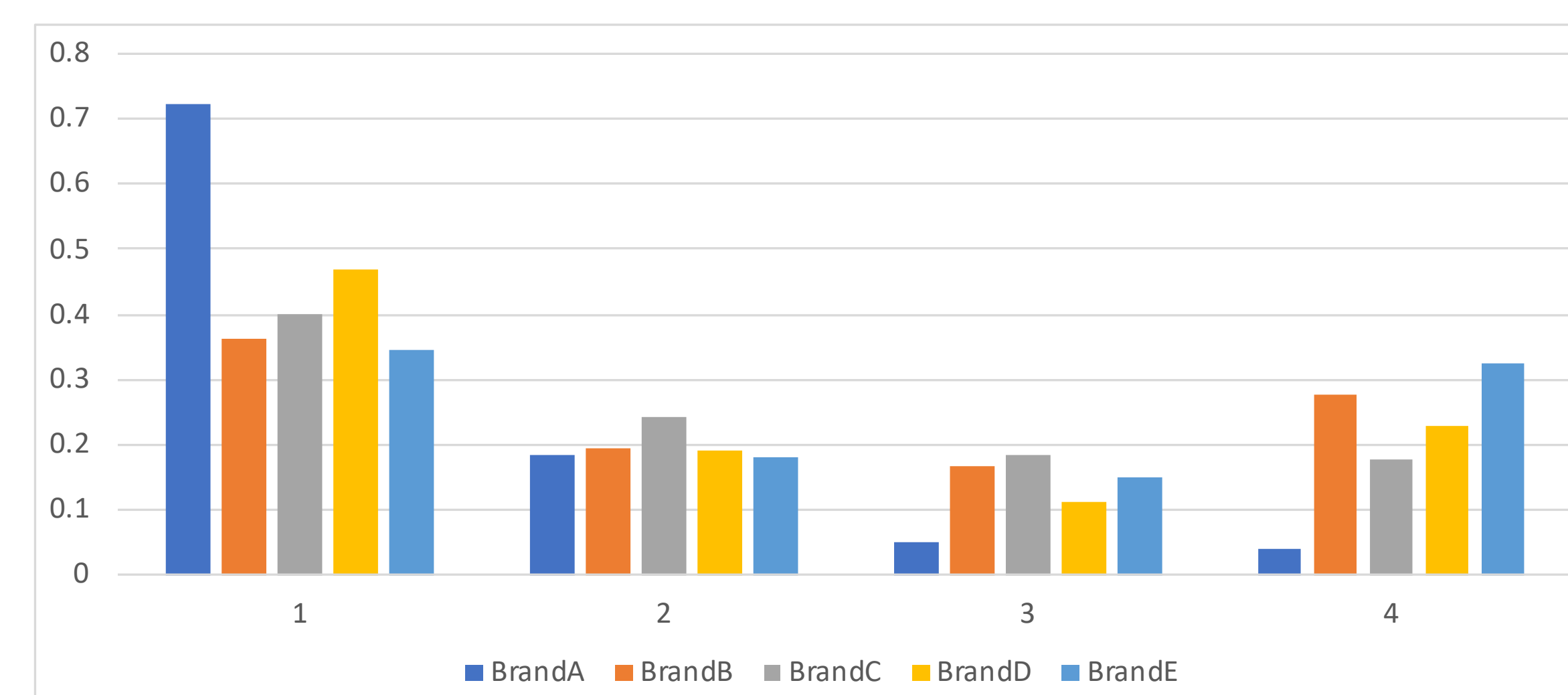
- Both proposed methods statistically significantly outperform the collections from experts
- For Bath, it is better to use domain knowledge to generate candidate sets while for Patio there is no difference in the two proposed methods

Discussion

Why collections from experts perform worse compared to the proposed?

01 Collections involve products that are interchangeable

02 Collections involve products of the same brand only, while consumers do not shop this way



- X axis: number of products that belong to the same brand
- Y axis: percentage of transactions

E.g.: For BrandA (blue bar) 72% of the transactions involved 1 product from the same brand BrandA (i.e., all other products were of different brand)