CASE STUDY

Building a recommendation system

Tuan Son Nguyen

Problems

- As we have all seen that commercial websites show related products through the lens of “customers who viewed also viewed” or “customers who purchased also purchased”.
- The objective of the case study is to identify related products that we can detect and show how we would detect them automatically for a e-commercial business
- This case study presents a recommendation model using Jaccard similarity.

E-Corp Data Set

	order_number	l1	l2	l3	sku	brand
0	168266	Power Tools	Power Saws and Accessories	Reciprocating Saw Blades	265105	2768
1	123986	Safety	Spill Control Supplies	Temporary Leak Repair	215839	586
2	158978	Hardware	Door Hardware	Thresholds	284756	1793
3	449035	Electronics, Appliances, and Batteries	Batteries	Standard Batteries	12579	1231
4	781232	Motors	General Purpose AC Motors	General Purpose AC Motors	194681	2603
5	116599	Pneumatics	Pneumatic Tube Fittings	Pneumatic Push to Connect Tube Fittings	167757	3889
6	701116	Motors	General Purpose AC Motors	General Purpose AC Motors	310296	1068
7	555497	Motors	Motor Supplies	Capacitors	306732	1068
8	282317	Safety	Footwear and Footwear Accessories	Insoles	148549	2696
9	644437	Hand Tools	Sockets and Bits	Crowfoot Socket Wrenches	283869	3356

E-Corp Data Set

```
RangeIndex: 2107537 entries, 0 to 2107536
Data columns (total 6 columns):
 #   Column        Dtype
---  ------        -----
 0   order_number  int64
 1   l1            object
 2   l2            object
 3   l3            object
 4   sku           int64
 5   brand         int64
dtypes: int64(3), object(3)
memory usage: 96.5+ MB
```

sku: item IDs

Recommendation Approach

- 1. Collaborative filtering
  - Approach:
    - Directly recommend item IDs (SKUs)
  - Technique:
    - Item-based
    - User-based
- 2. Recommendation implementation
  - Predicting item names, l3 category
  - Item-item similarity
  - Jaccard similarity coefficient: Intersection (A, B)/Union (A,B)

Recommendation Examples

Input	Recommendation			
Item1	Item1	Item2	Item3	Item4
Floor Finishes and Sealers	Hand Sanitizer, Lotion, and Soap Dispensers	Toilet Paper Dispensers	Dust Mops	Surface and Air Deodorants

Recommendation Examples

	Item Name	Brand IDs
Inputs		
Item1	Steel-Toe Work Boots and Shoes	746
Item2	Electrical Tapes	3706
Item3	Hex Head Cap Screws	1793
Recommendation		
Item1	Steel-Toe Work Boots and Shoes	2890
Item2	Steel-Toe Work Boots and Shoes	3486
Item3	Electrical Tapes	9
Item4	Electrical Tapes	3706
Item5	Hex Head Cap Screws	1793
Item6	Lock Washers	1793

Recommendation Examples

	Item Name	Brand IDs
Inputs		
Item1	Wrist Supports and Wraps	934
Item2	Steel-Toe Work Boots and Shoes	4746
Item3	Snips, Shears And Scissors	858
Recommendation		
Item1	Coated Gloves	2676
Item2	Cut-Resistant Sleeves	4713
Item3	Steel-Toe Work Boots and Shoes	4315
Item4	Steel-Toe Work Boots and Shoes	2890
Item5	Combination Wrench Sets	1698
Item6	Socket Bits	4692

Challenges in Recommendation Model

- 1. Large data set
  - Memory
  - Performance
- 2. Many SKUs
  - The same items' names with different brand IDs

Solutions

- 1. Use appropriate technology
  - Database organisation
- 2. Create recommendation model that considers characteristics of the data
  - Recommend a similar product with different brands
  - Recommend a related product

THANK YOU