BERKELEY IEOR
INDUSTRIAL ENGINEERING & OPERATIONS RESEARCH

IBM Research

Dynamic Learning with Frequent New Product Launches: A Sequential Multinomial Logit Bandit Problem

Junyu Cao* Wei Sun+

*University of California, Berkeley
+IBM Thomas J. Watson Research Center

Motivation

- Companies frequently release new products to learn customers' rapidly changing preferences

- New products carry more risks compared to existing products:
  - No history
  - Some with lower revenue as they could be intentionally priced low to attract customers

- Key research question: How can a company quickly learn customers' preferences while mitigating the risks inherent in new products?

Our Contributions

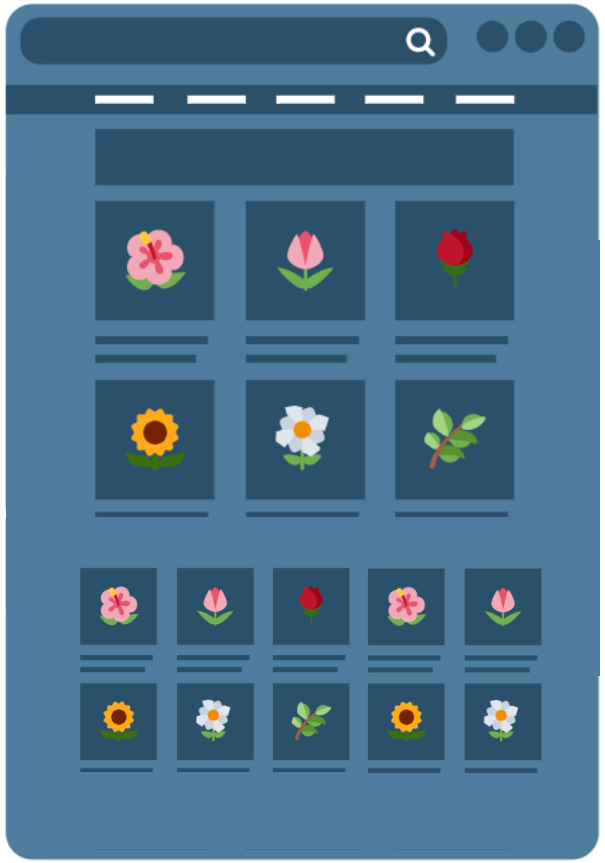
- We approach this problem as an online learning task
  - Seller's decision: which products to offer and how to display them
  - Seller's goal: maximize cumulative profit

- The setting is different from traditional literature
  - Frequent new product launches
  - Minimum learning criteria

- We show that a judicious choice of presenting products is capable of mitigating some costs associated with learning new products

Optimized Display of Multi-tiered Assortment

Landing page with 2-tier assortment



Primary assortment are displayed prominently to grab viewers' attention first, e.g., centrally positioned products with enlarged graphics, videos

Secondary assortment products are considered *after* primary assortment products

Probability of choosing product i from the first tier

$$p_i(\mathbf{S}) = \begin{cases} \dfrac{v_i}{1+\sum_{j\in S_1} v_j}, & \text{if } i \in S_1 \\ \dfrac{1}{1+\sum_{j\in S_1} v_j}\dfrac{v_i}{1+\sum_{j\in S_2} v_j}, & \text{if } i \in S_2 \\ \dfrac{1}{1+\sum_{j\in S_1} v_j}\dfrac{1}{1+\sum_{j\in S_2} v_j}, & \text{if } i = 0 \\ 0, & \text{otherwise,} \end{cases}$$

Probability of not choosing any products from the first tier

Online learning with Minimum Learning Criterion

- **Characterization of the optimal sequence:** Profit-ordered by tier
- **Minimum learning criterion:** Within ϵ accuracy with probability at least $1 - \delta$
- **Optimal placement strategy:** Display new product with low profit to the second tier
- **Regret analysis**

The regret during time $[0, T]$ is bounded above by

$$Reg_\pi(T; \mathbf{v}) \leq CK \log^2(KT) + C\sqrt{TK \log(KT)} + M \sum_{i \in X} v_i (r_{max} - r_i)$$

To find out more…

Poster Session:

Wed Jun 12th 06:30 -- 09:00 PM @ Pacific Ballroom #130