# A Debiased Machine Learning Framework for Optimizing Price Promotion in E-commerce

Bo Zhang (PayPal), Sergazy Nurbavliyev, Vaidyanath Areyur Shanthakumar, Stephen Merrill, Raymond J. Pan, Komson Chanprapan (Beyond Inc)

KDD Toronto 2025

# Problem

- Objective: Using data and model-driven algorithm to **recommend price promotion/discount levels** for millions of products

- Goal: Optimizing business metrics such as revenue or profit

- Constraint:
  - Budget X% of the total revenue
  - Discount levels [5%, 10%, 15%, 20%, 25%]

# Challenges with Existing Approaches

- Two step approaches:
  - Predict: Estimate individual level product demand sensitivity to different promotion levels
  - Optimization: Using the predicted sensitivities, apply optimization techniques to select the optimal discounts for each product by maximizing overall financial performance under business constraints.

- Issues:
  - In industry practice, the focus is often on improving demand prediction accuracy rather than estimating causal impacts.
  - Uplift modeling requires unbiased experimental data, which is often difficult and costly to obtain.

# Delta Method

- Inspired by fixed-effects regression in economics. The intution is that the impact is driven by within-product variation via de-meaning variables.

$$y_{it} - \overline{y}_i = (X_{it} - \overline{X}_i)\beta + u_{it} - \overline{u}_i$$

- Similary, we de-meaning the y and x in a machine learning model to get the individual treatment effects

$$qty_{it} - \overline{qty}_i = f(dsc_{it} - \overline{dsc_i}, X_{it} - \overline{X}_i, X_{it})$$

# Evaluation: Aggregated Impact

**Table 2: LightGBM: Standard vs Delta**

| | sample average | standard LightGBM | delta LightGBM |
|---|---|---|---|
| | (1) | (2) | (3) |
| dsc=0% | 0.084 | 0.266 | 0.224 |
| dsc=5% | 0.543 | 0.372 | 0.230 |
| dsc=10% | 0.418 | 0.299 | 0.246 |
| dsc=15% | 0.448 | 0.313 | 0.286 |
| dsc=20% | 0.308 | 0.202 | 0.320 |
| dsc=25% | 0.199 | 0.202 | 0.346 |

*Note*: Column (1) shows the average product quantity sold at different discounts.

**Table 3: Neural network: Standard vs Delta**

| | sample average | standard network | delta network |
|---|---|---|---|
| | (1) | (2) | (3) |
| dsc=0% | 0.084 | 0.164 | 0.099 |
| dsc=5% | 0.543 | 0.227 | 0.180 |
| dsc=10% | 0.418 | 0.260 | 0.260 |
| dsc=15% | 0.448 | 0.274 | 0.330 |
| dsc=20% | 0.308 | 0.316 | 0.393 |
| dsc=25% | 0.199 | 0.294 | 0.451 |

*Note*: Column (1) shows the average product quantity sold at different discounts.

# Evaluation: Individual Effect Using Standard Method

**Figure 2: The quantitative relationship between discount levels and product quantity sold for 5 randomly selected products from the standard LightGBM**

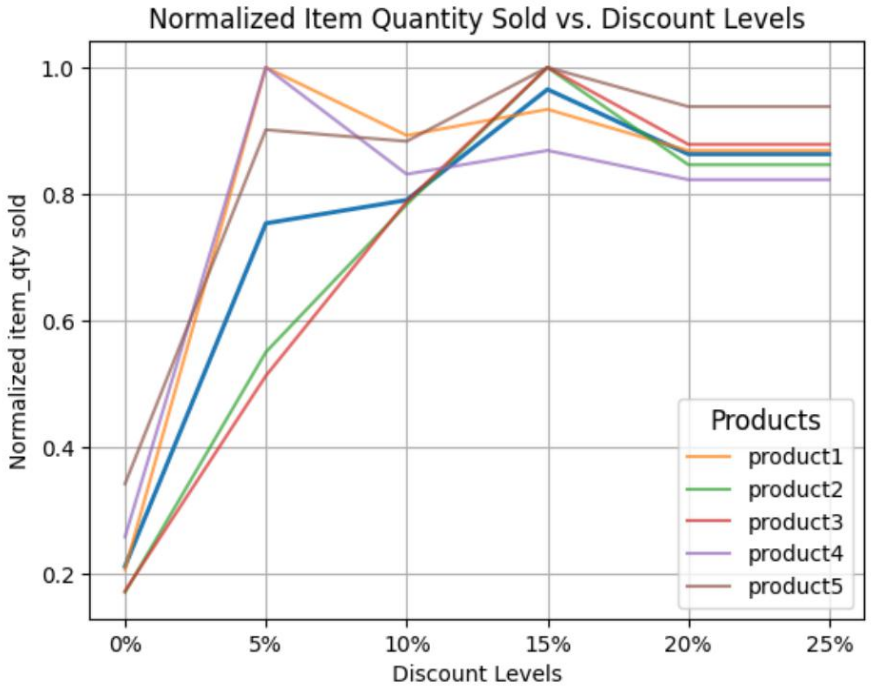
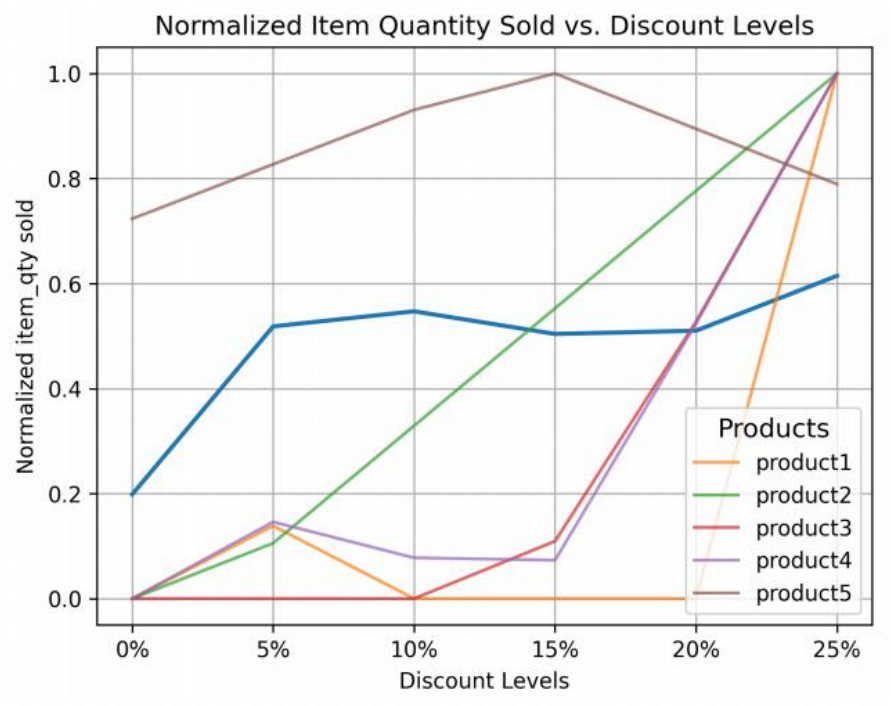



**Figure 4: The quantitative relationship between discount levels and product quantity sold for 5 randomly selected products from the standard neural network model**



# Evaluation: Individual Effect Using Delta method

**Figure 3: The quantitative relationship between discount levels and product quantity sold for 5 randomly selected products from the LightGBM with Delta Method**

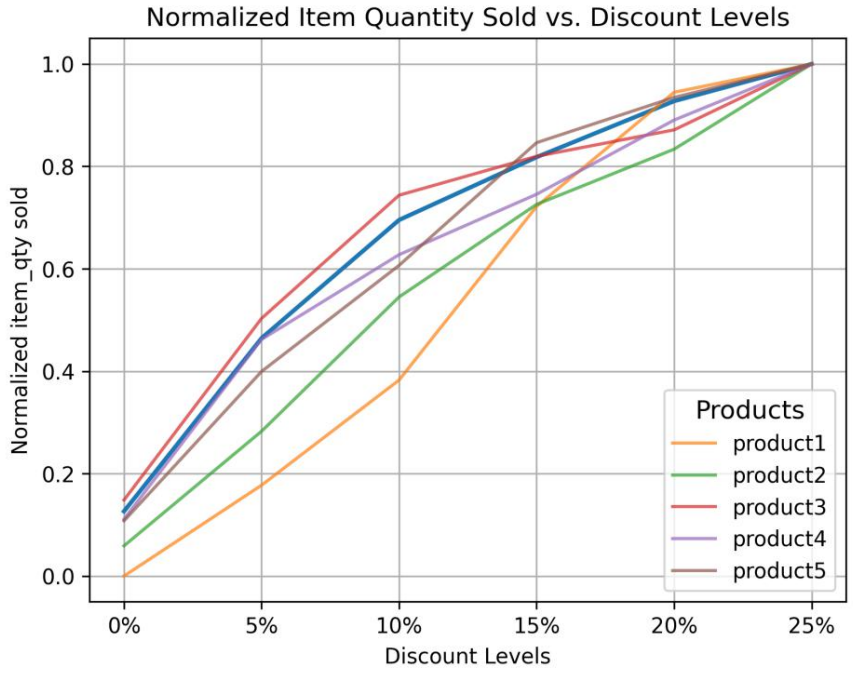
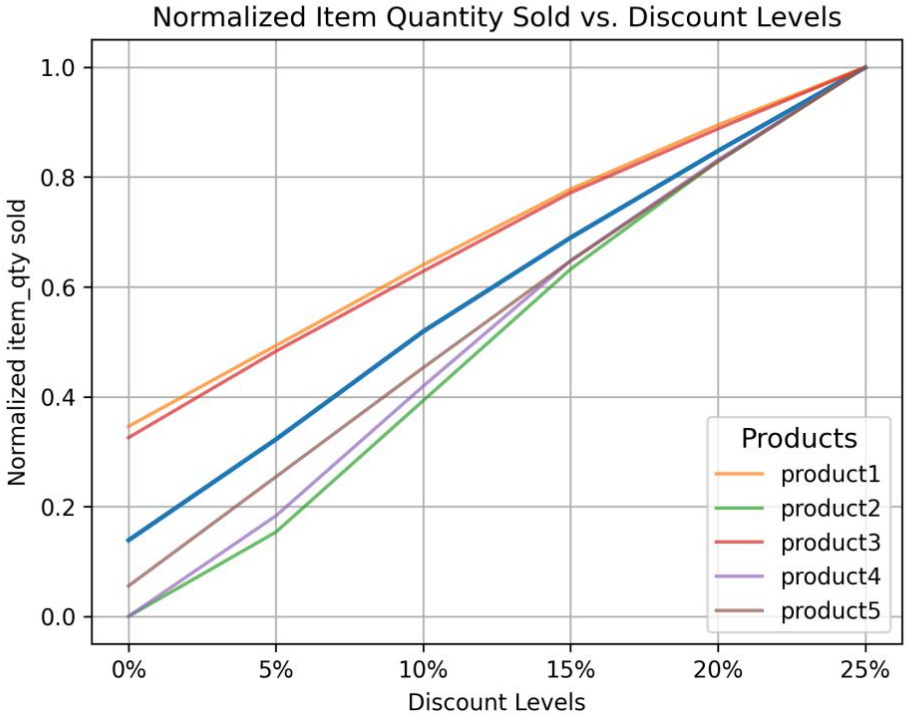



**Figure 5: The quantitative relationship between discount levels and product quantity sold for 5 randomly selected products from the neural network model with Delta Method**



# Evaluation: Accuracy

**Table 4: Standard method model performance**

| | Regression | LightGBM | Neural Network |
|---|---|---|---|
| MAPE | 1.914 | 0.852 | 1.116 |
| weighted_MAPE | 0.847 | 0.529 | 0.827 |
| Person Correlation | 0.230 | 0.757 | 0.195 |
| MAE | 6.872 | 3.875 | 3.927 |
| MAD | 4.610 | 1.605 | 2.032 |
| R2 | 0.044 | 0.622 | 0.570 |

**Table 5: Delta method model performance**

| | Regression | LightGBM | Neural Network |
|---|---|---|---|
| MAPE | 1.188 | 0.836 | 0.798 |
| weighted_MAPE | 1.249 | 0.502 | 0.511 |
| Person Correlation | 0.636 | 0.840 | 0.831 |
| MAE | 5.508 | 3.712 | 3.687 |
| MAD | 2.371 | 1.554 | 1.549 |
| R2 | 0.307 | 0.703 | 0.690 |

# Results — Real-World Impact

- A/B Test: 3 million products randomized, 2-week experiment.
- Delta Method vs. standard algorithm:
- Revenue: +3%
- Profit: +2%
- Statistically significant improvements, robust to business constraints.

# Q&A