Predicting Product Popularity

With Amazon dataset

Brain Station Capstone

Soniya Sharma

Understanding product popularity

- Millions of products →
  Overwhelming choices for buyers

- Ratings shape consumer trust and influence purchasing decisions.
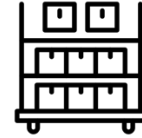
- Can we predict success on amazon?

The Potential Impact

Consumers

✓ Better purchasing decisions

✓ Fewer Returns

Retailers

✓ Optimized Inventory

✓ Refined Marketing

E-Commerce Platforms

✓ Enhanced Trust

✓ Higher Engagement

Overview of Datasets

Meta Data

- 164k products
- 12 main product categories
- > 3800 sellers
- Product prices
- Product images
- Product features (material/brand/dimensions etc.)

Review Data

- 664k reviews
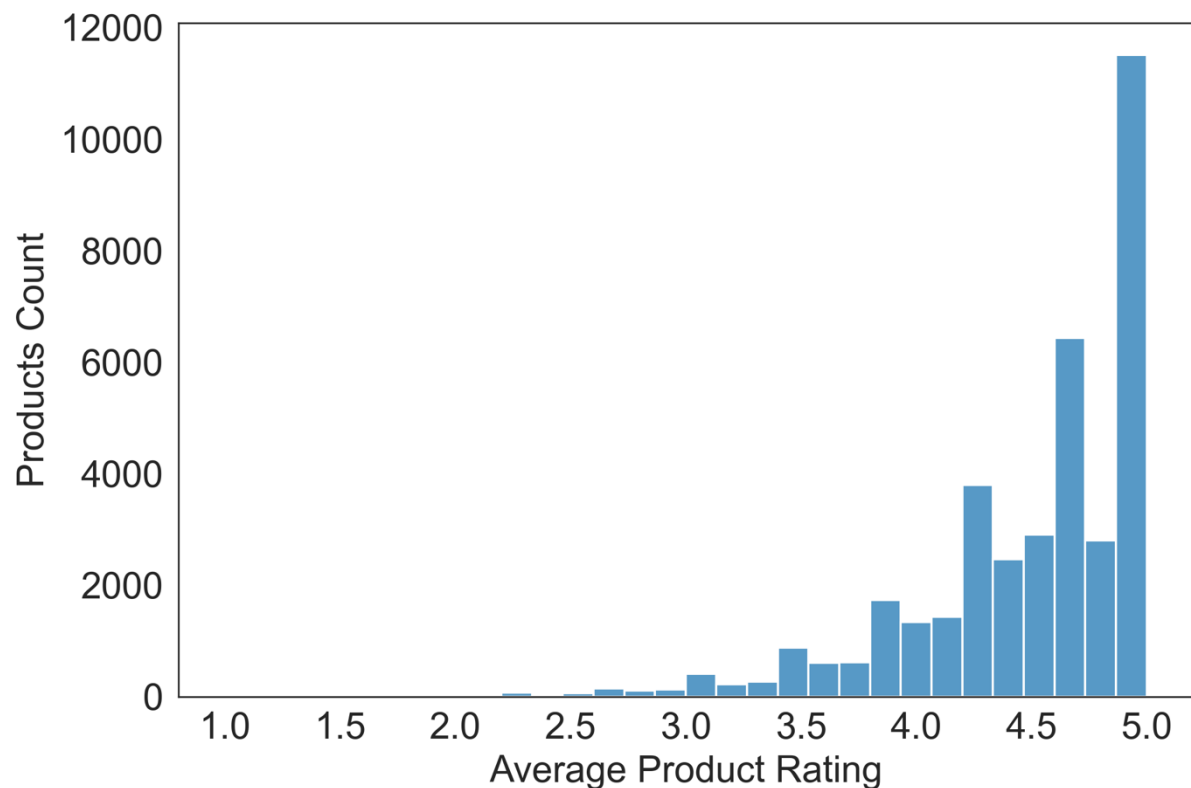- Verified purchases
- Review text
- Helpful votes

Aggregated Metrics per Seller/Category

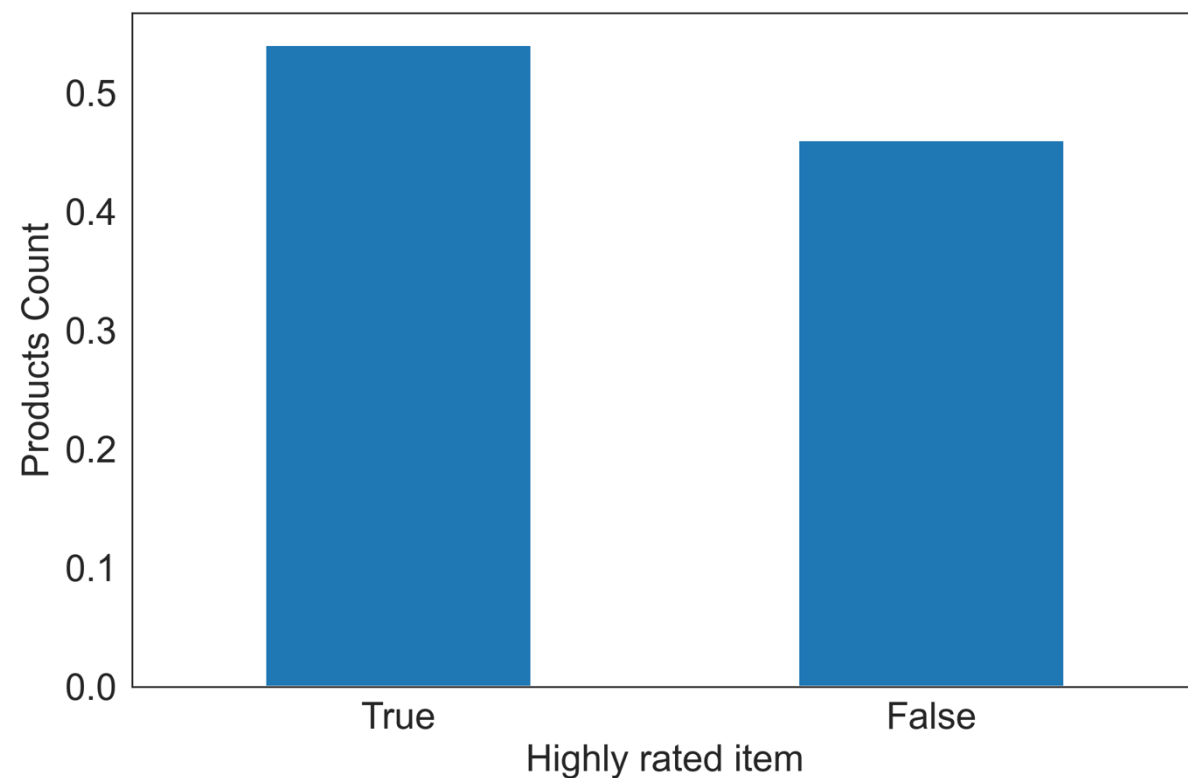*Data Source: McAuley Lab, collected in 2023 on the category of handmade items.*

From Ratings to Target Variable

Target variable shows slight imbalance, with most products receiving high ratings.
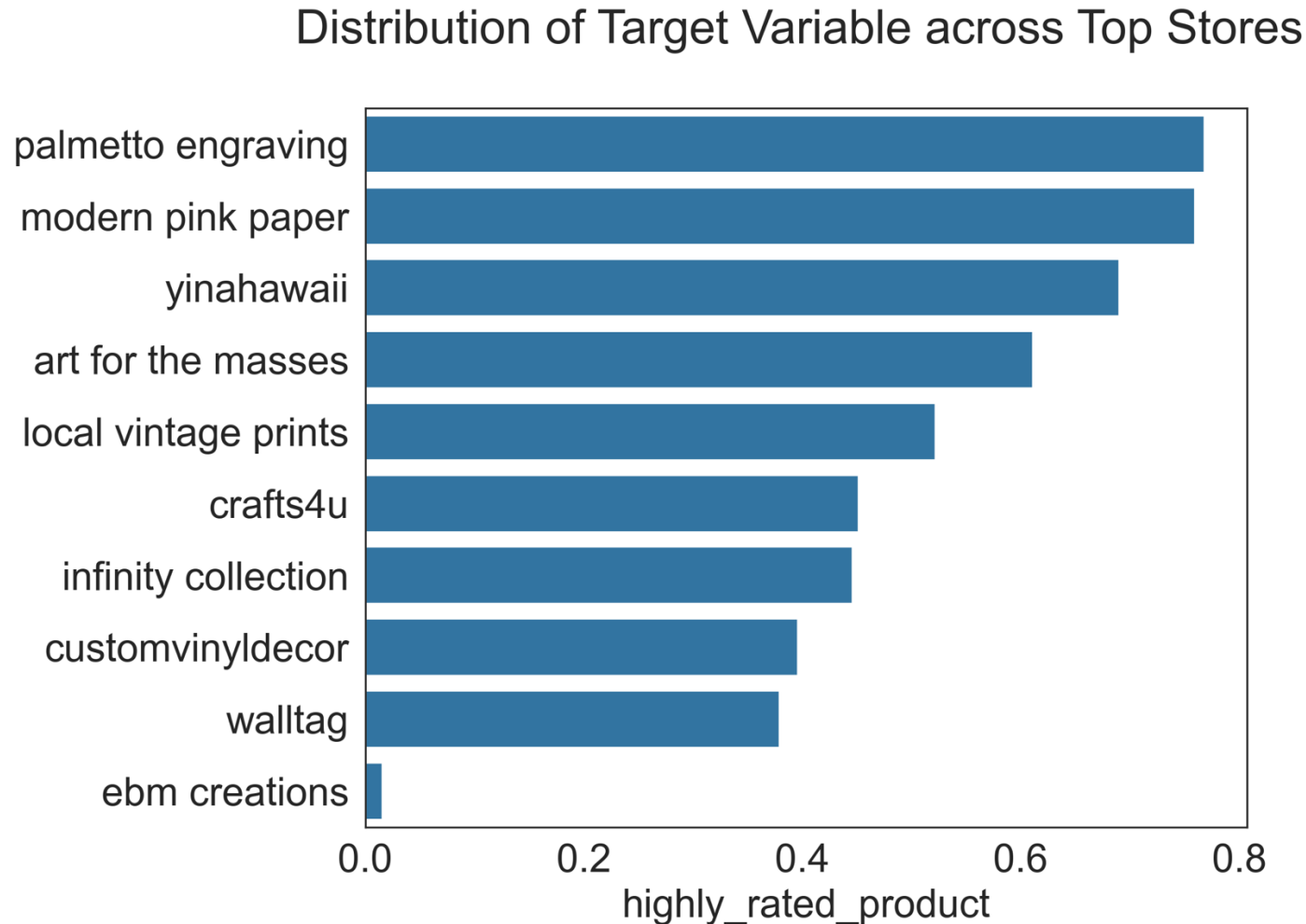
What drives Product Ratings?

Seller reputation plays a significant role in product success.

Distribution of Target Variable across Top Stores

Building a Baseline: Logistic Regression

ACCURACY	64%
PRECISION (class 0)	64%
RECALL (class 0)	48%
ROC-AUC SCORE	69%
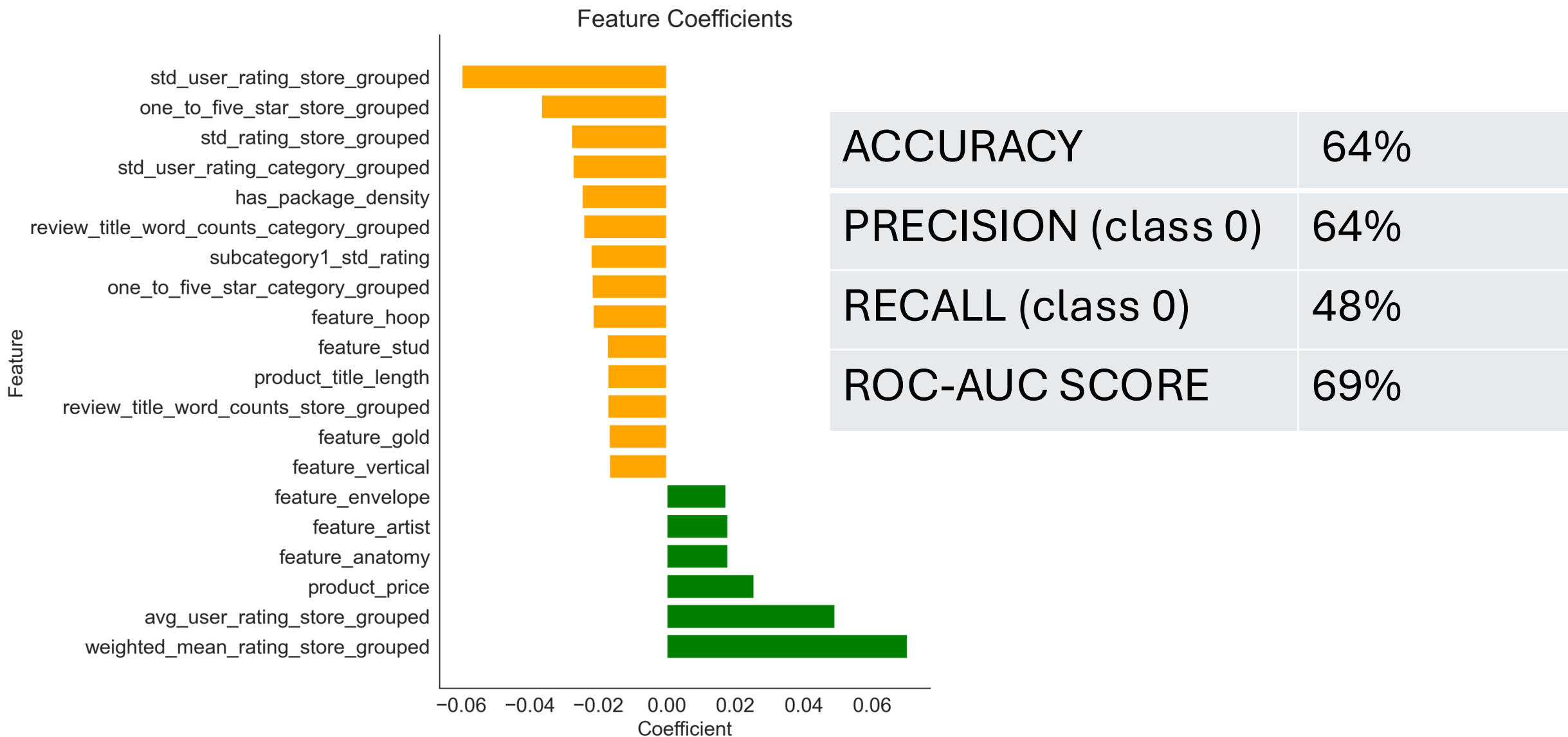
Beyond Popularity Prediction– Understanding Consumer Choices

Offer Personalized Recommendations based on user preferences and behavior.