



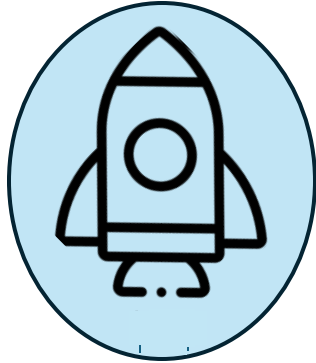
Predicting Product Popularity

With Amazon dataset

Brain Station Capstone
Soniya Sharma

THE PROBLEM

Companies want to gauge product reception before launch to optimize strategy.



Pre-launch Prediction

Meta-data (e.g., Features/ Price/Category)



Can we predict product reception before launch?

Early Insights



Post-launch Feedback

- ✓ Sales data
- ✓ Customer Reviews



Model Validation

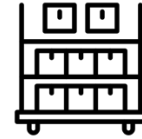
THE POTENTIAL IMPACT

Consumers



- ✓ Better purchasing decisions
- ✓ Fewer Returns

Retailers



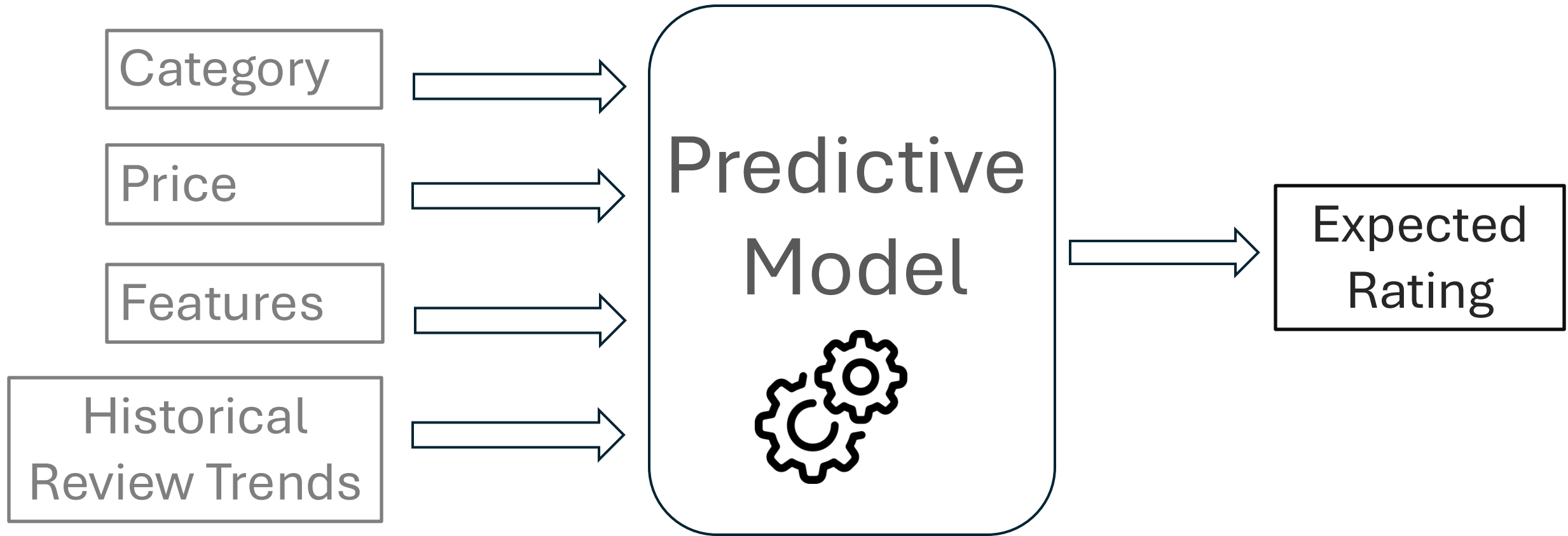
- ✓ Optimized Inventory
- ✓ Refined Marketing

E-Commerce Platforms



- ✓ Enhanced Trust
- ✓ Higher Engagement

THE PROPOSED VISION



Challenges



Extracting insights from
meta-data



Review Quality
Variability



Data Imbalance

THE AMAZON DATASET By McAuley Lab

Main Category:
Handmade Items

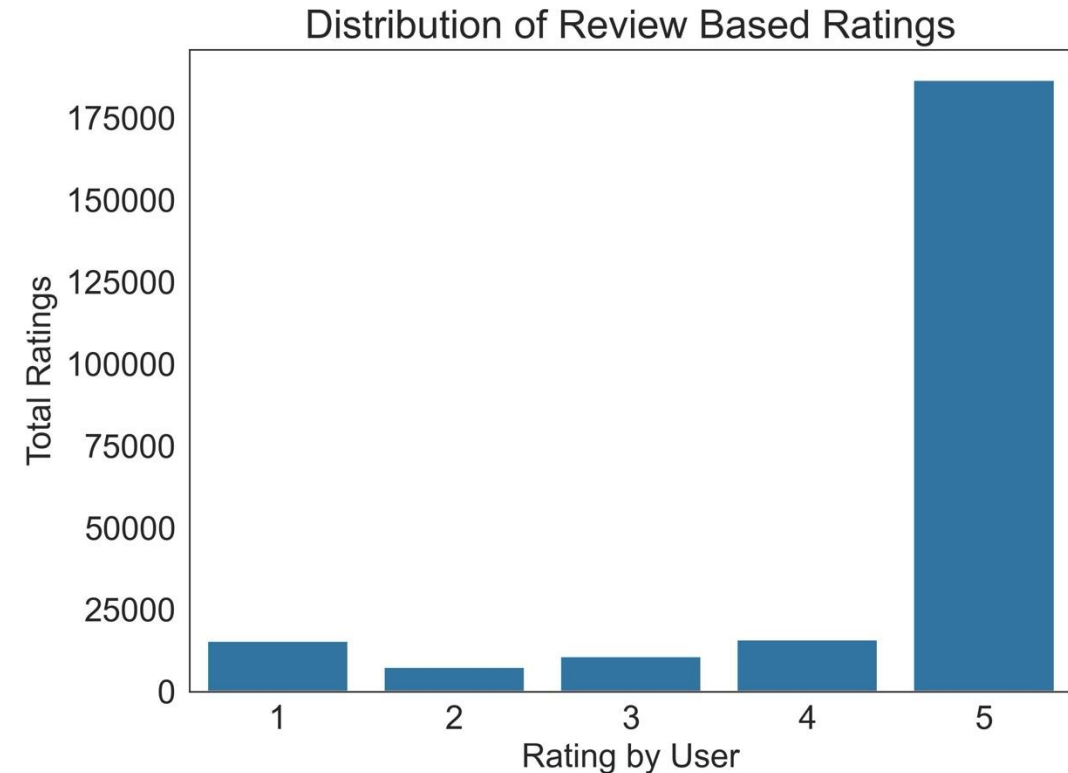
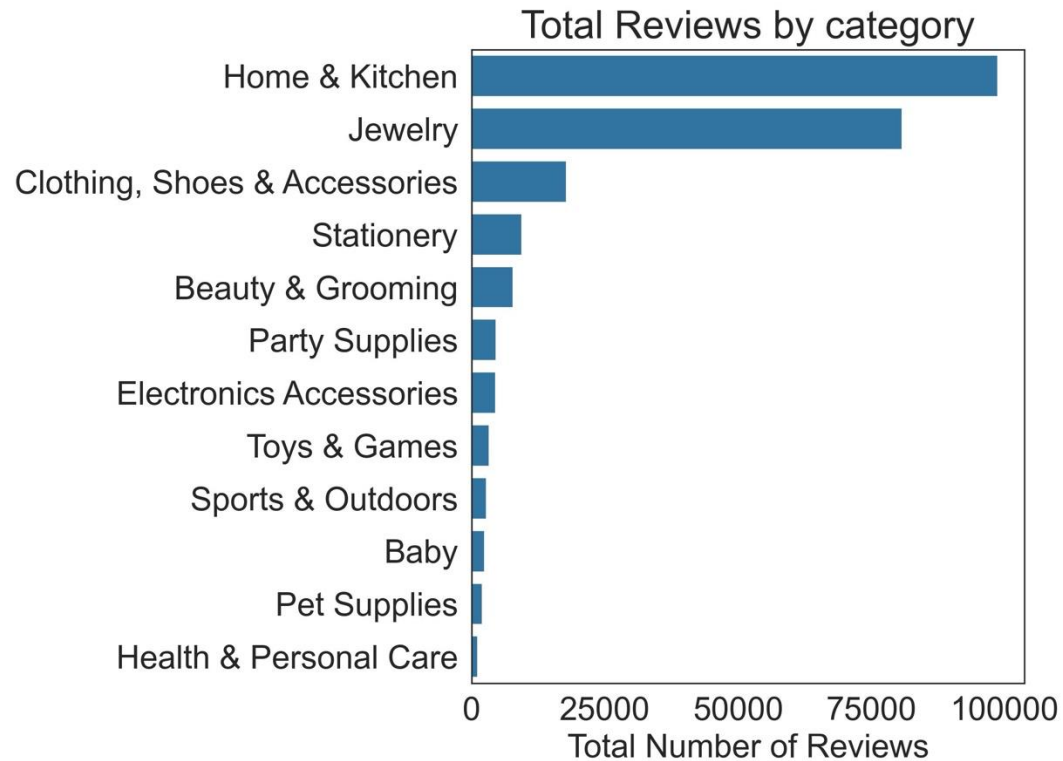
Meta Data

- 164k products
- 500 Hierarchical product categories



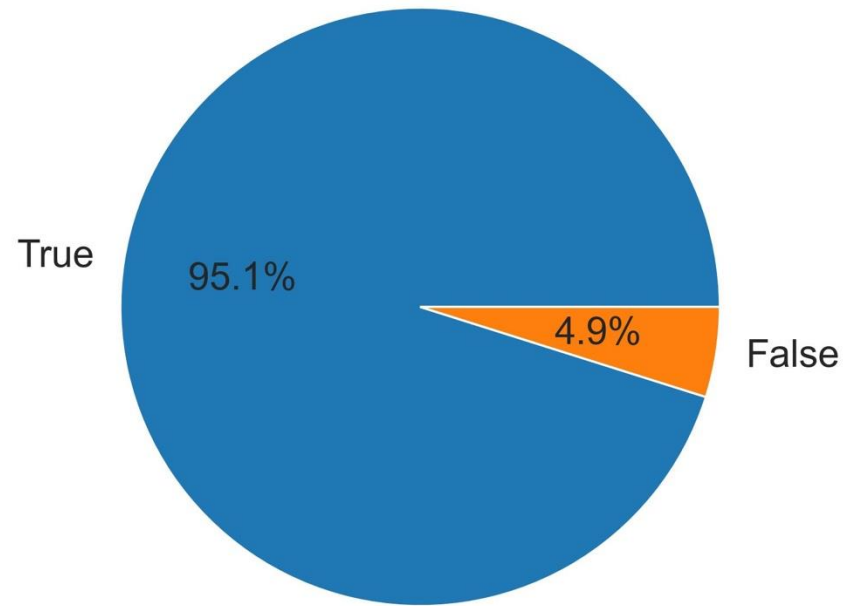
Review Data

- 664k reviews
- Verified purchases
- Helpful votes

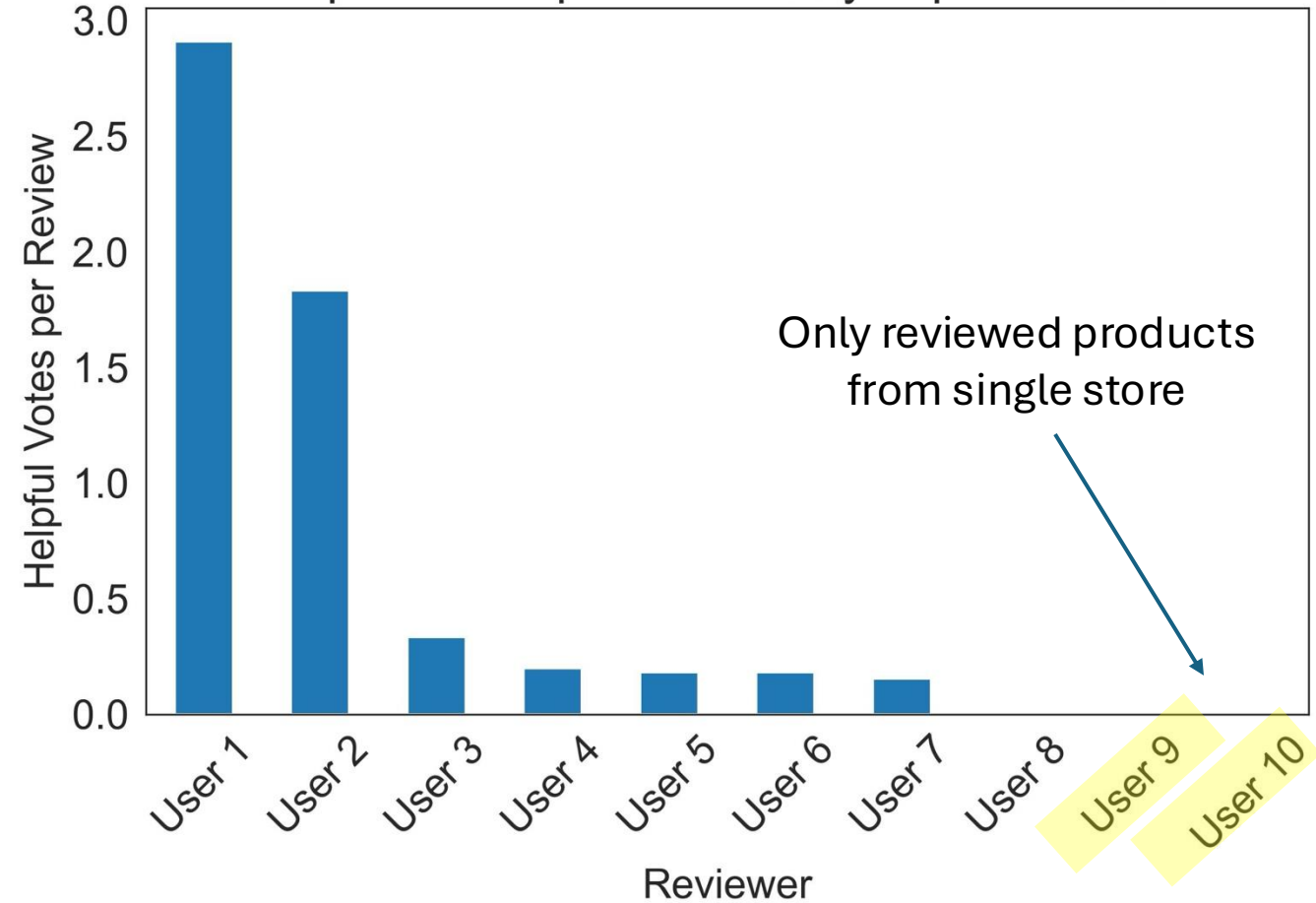


CAN ALL REVIEWS BE TRUSTED?

Percentage of verified vs not verified purchases



Helpful Votes per Review by Top Reviewers



NEXT STEPS IN ANALYSIS

Text Processing

Feature Engineering

Baseline Modeling

- Sentiment Analysis of Reviews
- Insights from product title and descriptions
- Dummies for product categories with significant items
- Fit Classification model to predict whether an item is highly rated or not

