

Trending Products: Amazon Toy Reviews

The next great review!

Jason Salazer-Adams

Trending Products: Amazon Toy Reviews

The next great review!

The next **terrific** review!

Jason Salazer-Adams

Overview

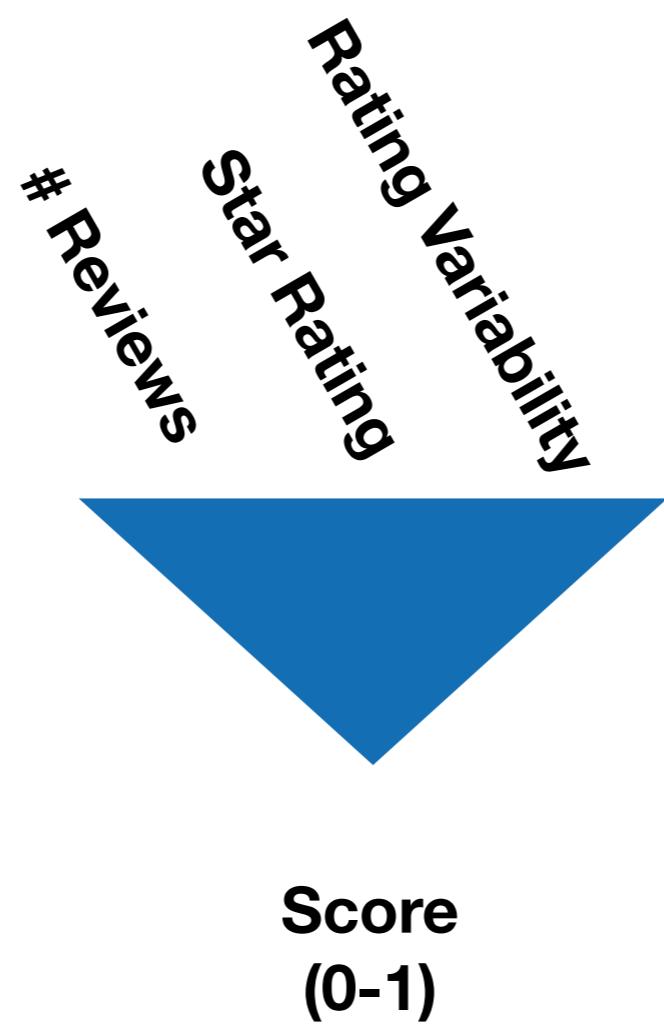


Process

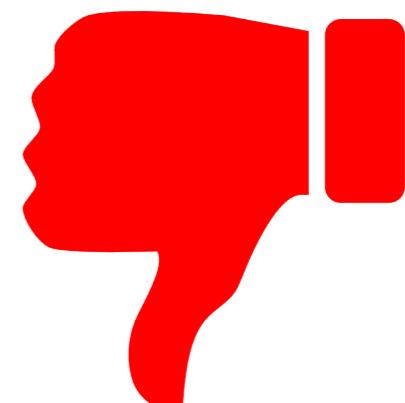
Review data 2014 - 2015



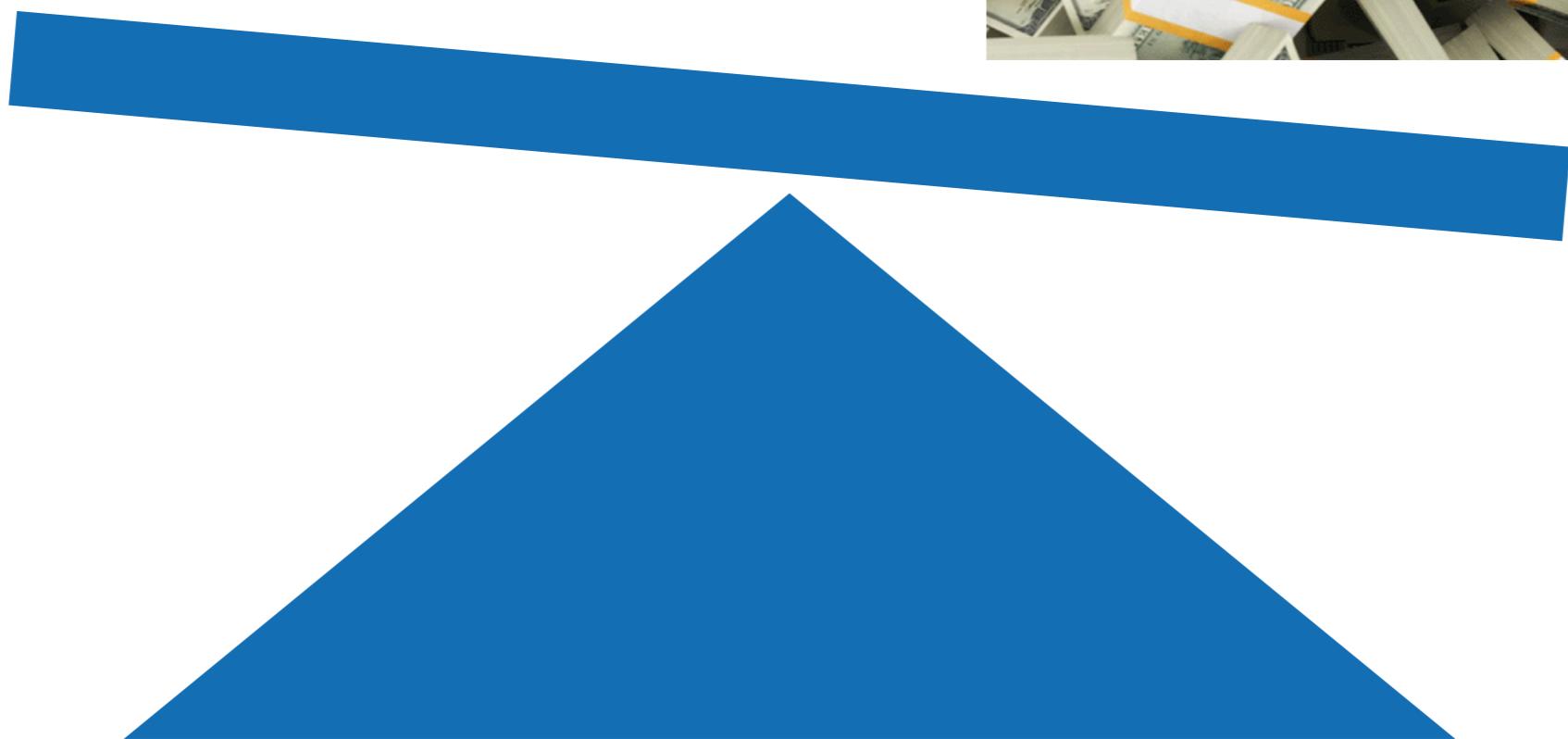
Trend Score



Trend or Not



Results



Results



66%

Review Improvement

“My family loves this game. Lots of non electronic time.”

“The next great review!”

Can a review be better?

Amazon trending product review generator

Please type your review here

Submit

Original
Improved

Trend Probability

Future Work

- Add more review history
- Improve review generation to be phrase replacement

Thank you

jandlsalazeradams@gmail.com

linkedin.com/in/jason-salazer-adams

github.com/jason-sa

Appendix

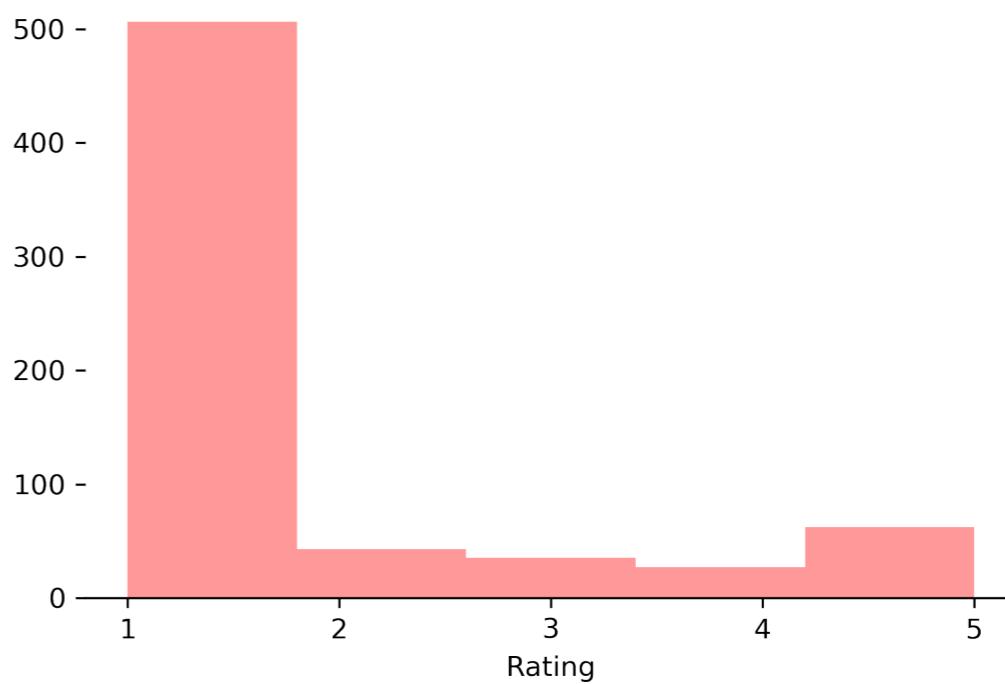
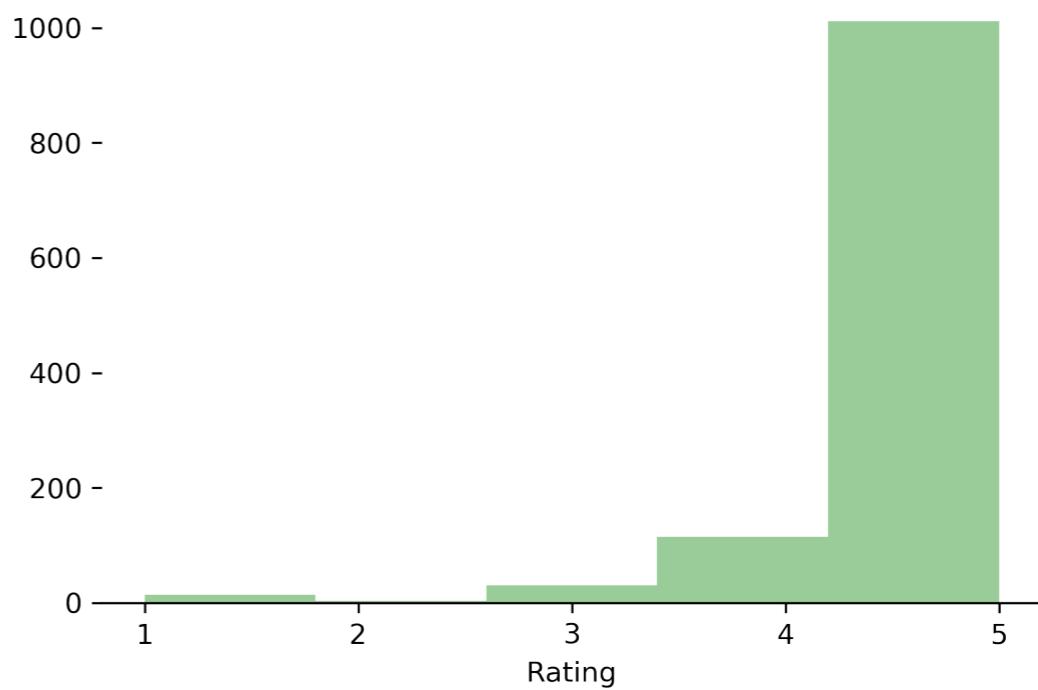
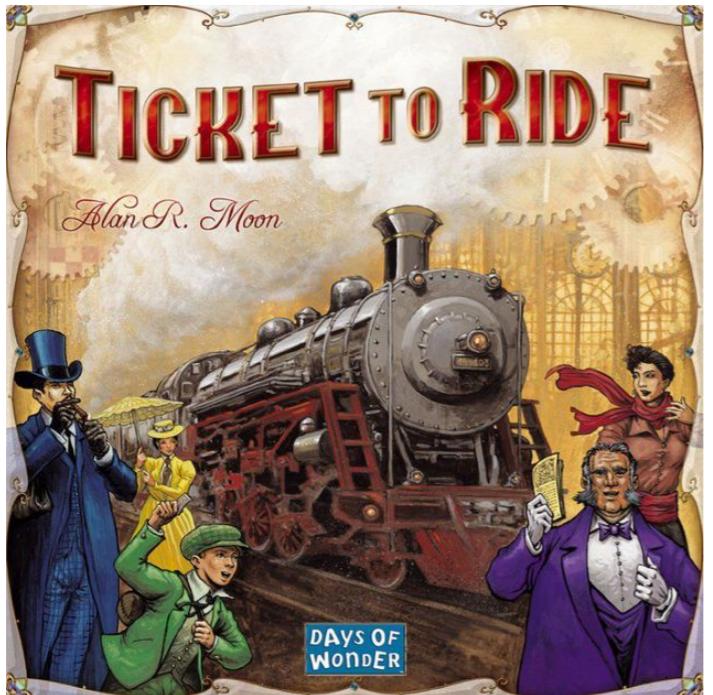
Review Characteristics

Trending Products

Not Trending Products

Trending Product

Trending product: Lots of reviews + High Ratings + Small variability



Trend Score Calculation

- Review Rate (RR) = Total Reviews / 30 (window in days)
- Adj Star Rating (ASR) = Star Rating $\wedge 1.5$
- Review Success (RS) = $(RR * \text{Median ASR}) / \text{Std ASR}$
- Trend Score = $\tanh(RS)$
- Top 1% of trend score is identified as trending

Model Selection - AUC

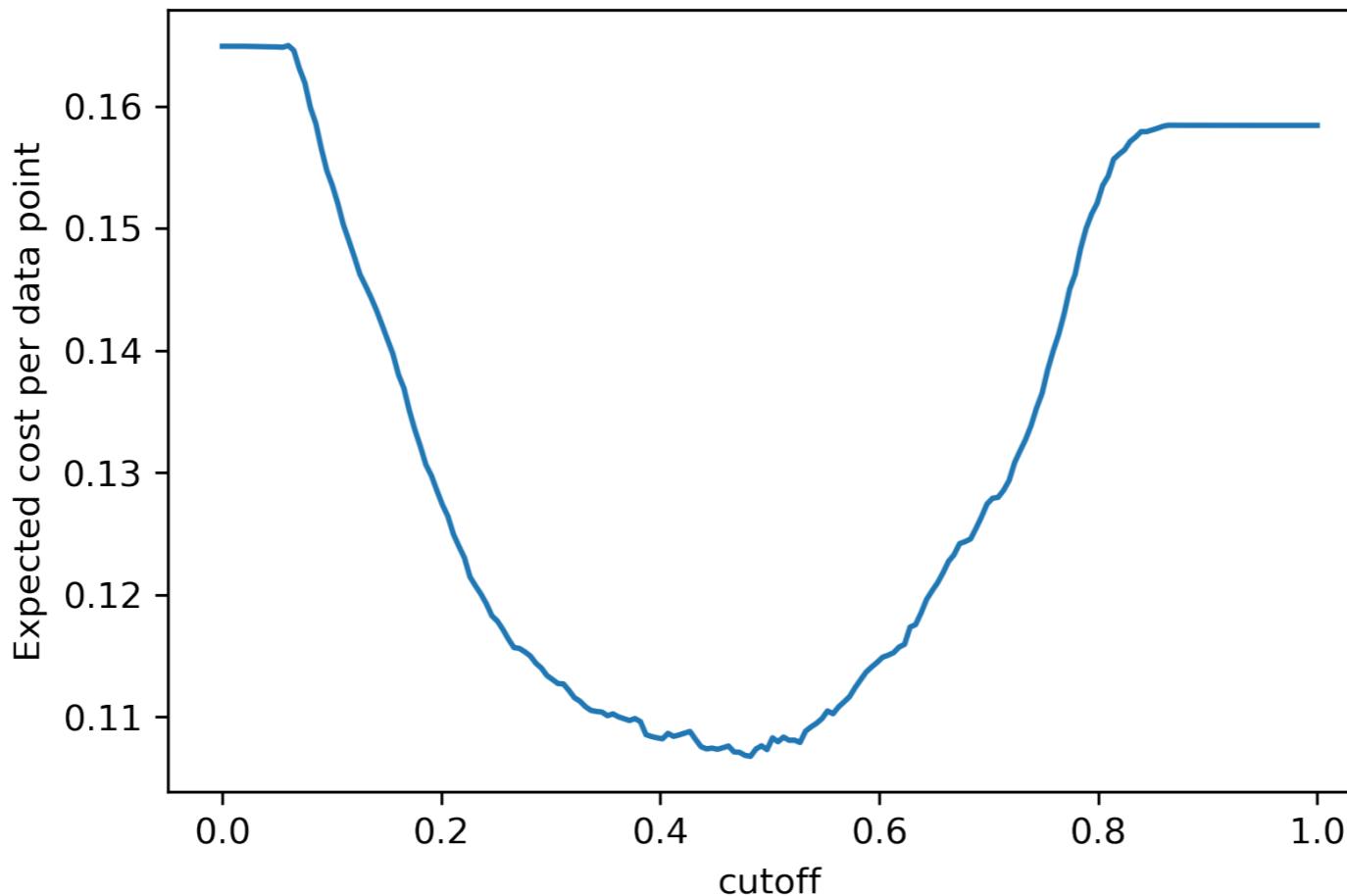
- Only utilized XGBoost trees.
- Performed 5-fold CV with BayesSearch to tune,
 - max_depth
 - learning_rate
 - n_estimators
 - gamma

| | Validation AUC |
|------------------------|----------------|
| 5 topic | 0.711 |
| 5 topic SMOTE | 0.712 |
| 10 topic | 0.711 |
| 10 topic SOMTE + Tomek | 0.716 |

5 topic SMOTE was chosen for simplicity, and the problem is imbalanced.

Cost optimization

- Let's assume we have a generic toy with COGS of \$10, a standard monthly inventory holding rate of 1.6%, and a markup of 100%. The error costs would be calculated as,
 - False positive cost = COGS * holding = $10 * .016 = \$0.16$
 - False negative cost = COGS * markup = $10 * 1.5 = 15$



Optimal Probability Cutoff: 0.48

Test Results - Confusion Matrix

| | Predicted No | Predicted Yes |
|------------|--------------|---------------|
| Actual No | 33480 | 18762 |
| Actual Yes | 168 | 390 |