

Enhance Amazon's Product Recommendation, Search, and Reviews

Sue Lim
February 5, 2023



BUSINESS & PROJECT OVERVIEW

Business Overview

- **Product managers** are looking to improve user experience
- User experience consists of the following:
 - Customized **recommendations**
 - Accurate **item search**
 - **Reviews** useful for decision-making

Enhance Amazon Product Recommendation, Search, and Reviews

- **Product Recommendation**
 - Model-based algorithms
- **Product Search**
 - Identify similar groups of products & expand search
- **Product Reviews**
 - Identify key sentences in positive/negative reviews which would be useful to customers

DATA

Amazon Product Reviews from UCSD

- Category “**Pet Supplies**” for years 2013 – 2018
 - 200K items, 3M users, and 6M reviews & ratings (average is 4 stars)
 - Subset down to 8K items, 3K users, and 90K reviews & ratings for speed of work

Amazon Berkeley Objects ("ABO") data

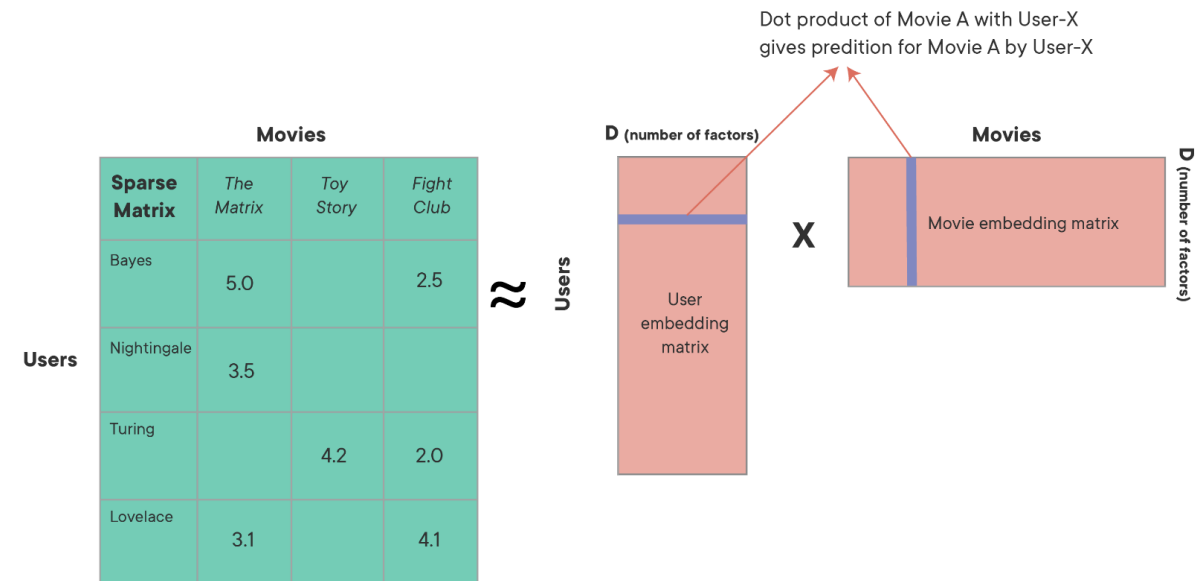
- The data contains **images**, **tags**, and other information for 50K products of various categories
- This project focuses on the category "**Shoes**" with 2.5K products



I. PRODUCT RECOMMENDATION

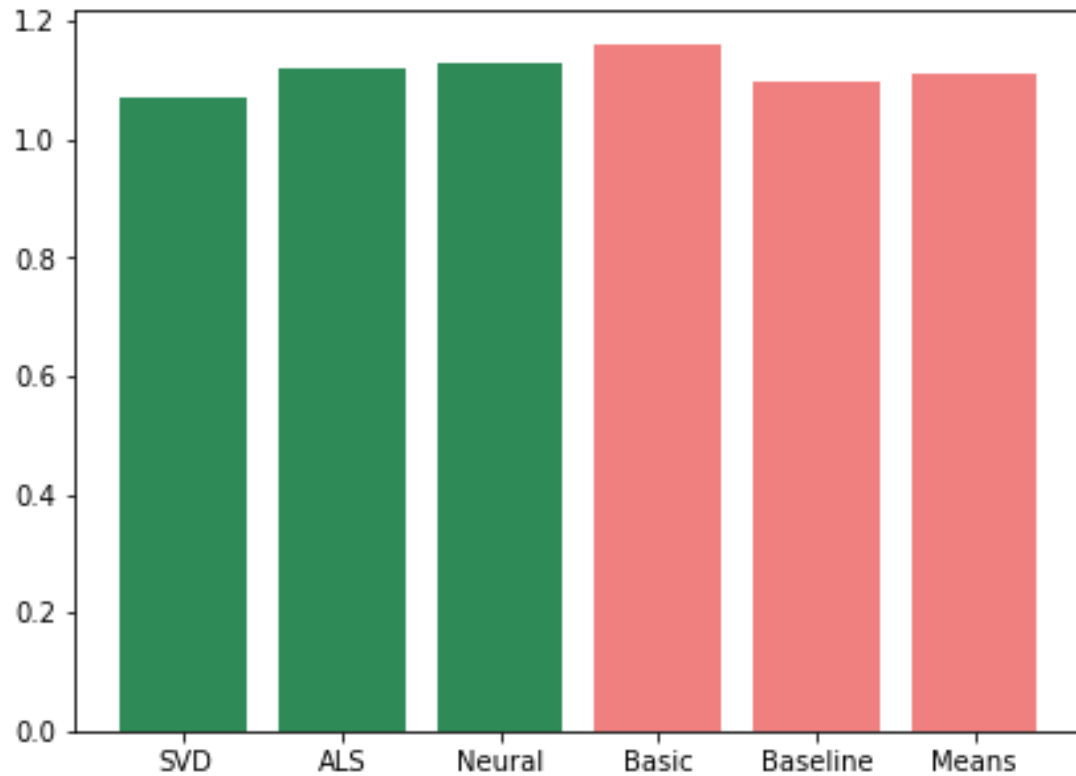
Enhance Product Recommendation

- **Model-based** approaches
- Compared to memory-based:
 - Scales better
 - Mitigates cold-start problems

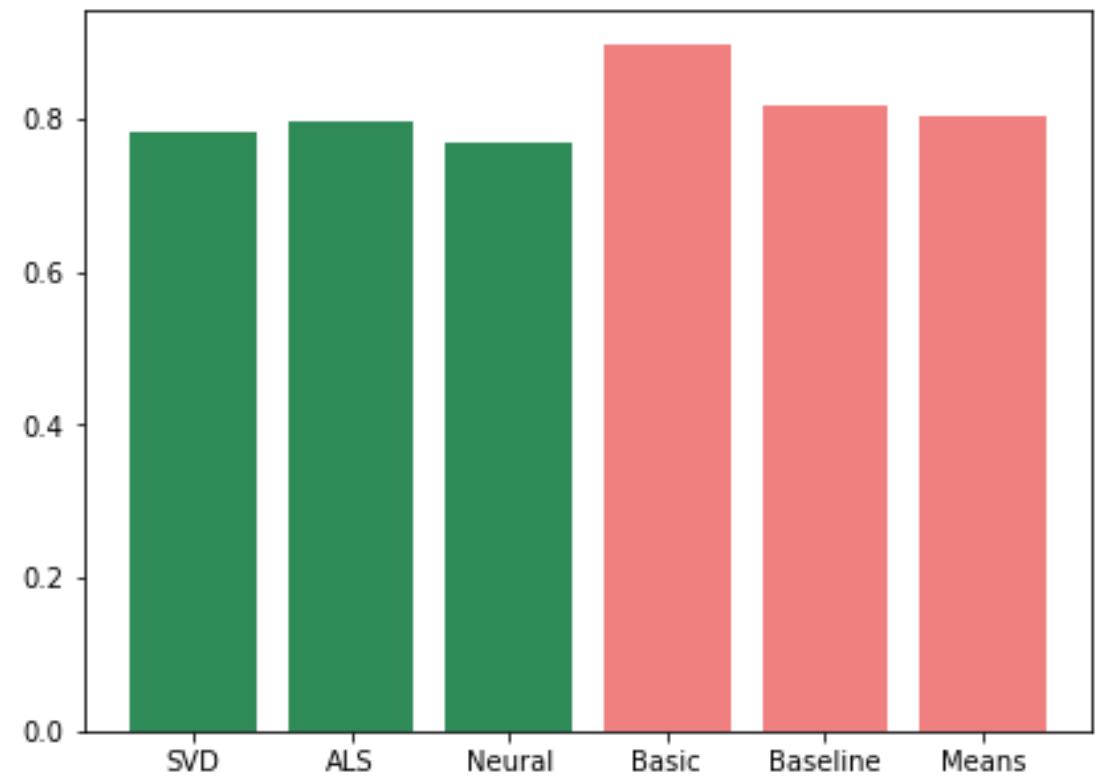


Comparison of Prediction Accuracy

Test RMSE



Test MAE



II. PRODUCT SEARCH

Enhance Product Search

- Identify similar groups of products & enable expanded search based on the categories.
- Method: **Multi-label classification** neural network
 - Step 1) Generate categories using keywords sellers post



Categories

- Modern sofa
- Ivory sofa

- Step 2) Use **product images** to predict **categories**

Example: Keyword “Formal Shoes”

- 4 categories containing **formal**
 - **formal** shoes for men black
 - **formal** shoes for mens leather
 - leather shoes for men **formal** branded
 - shoes for men **formal**



III. PRODUCT REVIEWS

Product Reviews

- Identify **key sentences** from **most popular reviews**
- Method: Term Frequency-Inverse Document Frequency (TF-IDF)
 - Step 1) Identify keywords from **positive/negative** reviews
 - Step 2) **Extract sentences** containing such keywords from most popular reviews

Example: “Nylabone Dura Chew Textured Dog Chew, X-Large”



Example: Top 5 Positive Reviews

- “(2560+ lbs)\nIt's really cute to walk along the aisle and see contented dogs happily holding their chew bones in their paws and gnawing away.”
- “I am so glad I got it and I suspect they are even happier!”
- **“My 60 pound boxer pit mix is a fan.”**
- **“Yummy & Healthy & Fun ...”**
- **“Wears slowly.”**

Example: Top 5 Negative Reviews

- **“Like I handed her a brick.”**
- “I was told that if we got her something like this, she would not tear up anything, like my Bible, anymore.”
- **“Price is very high than local store**, you may able to buy it from Marshall or other local store with better price, and my dog evening bleeding after play a while with this product, after one time use, I just through it away.”
- “so this is a big fat nope is our book of chew toys.”
- “I bought this when I had 4 dogs in the house (our two, and two puppies we were fostering), **out of 4 dogs NONE of them wanted this!**”

CONCLUSION & FUTURE WORK

Enhance Amazon Product Recommendation, Search, and Reviews

- **Product Recommendation**
 - Model-based algorithms
- **Product Search**
 - Identify similar groups of products & expand search
- **Product Reviews**
 - Identify key sentences in positive/negative reviews which would be useful to customers

And in the Future...

- **Product Recommendation**
 - Potentially explore hybrid approaches for best performance
- **Product Search**
 - Further validate the accuracy of expanded item search
- **Product Reviews**
 - Try different formulas for pulling keywords to see which extract most useful sentences
 - Produce an evaluation tool to measure the success