AMAZON PRODUCT

RECOMMENDATION SYSTEMS

By Natasha Gitlin

January 27, 2024

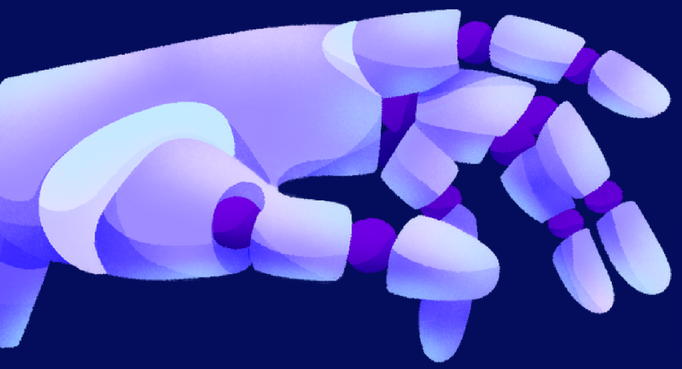
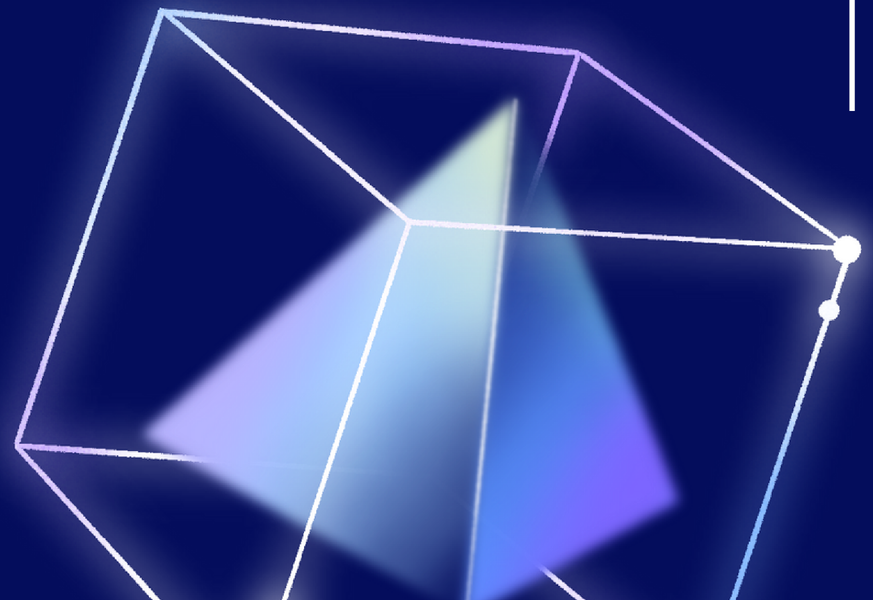


TABLE OF CONTENTS



BUSINESS PROBLEM

The rapid growth of information has resulted in overload and a paradox of choice for consumers. **Recommender Systems**, used by companies like Amazon, provide personalized product suggestions during online browsing, enhancing user engagement and business. Amazon's algorithm, such as item-to-item collaborative filtering, exemplifies AI's role in real-time, high-quality recommendations.

DATA OVERVIEW

- **userId:** Unique identifier for each user
- **productId:** Unique identifier for each product
- **Rating:** User-assigned rating for the corresponding product
- **timestamp:** Time of the rating (not used for the current task)

The dataset aims to utilize this information to build a recommendation system, emphasizing user-product interactions based on ratings for personalized suggestions.

EXPLORATORY DATA ANALYSIS

*The dataset was very large, with 7,824,482 observations so had to be reduced

No. of rows: 65,290

No. of columns: 3

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 65290 entries, 1310 to 7824427
Data columns (total 3 columns):
 #   Column   Non-Null Count  Dtype  
---  ------   --------------  -----  
 0   user_id  65290 non-null  object 
 1   prod_id  65290 non-null  object 
 2   rating   65290 non-null  float64
dtypes: float64(1), object(2)
memory usage: 2.0+ MB
```

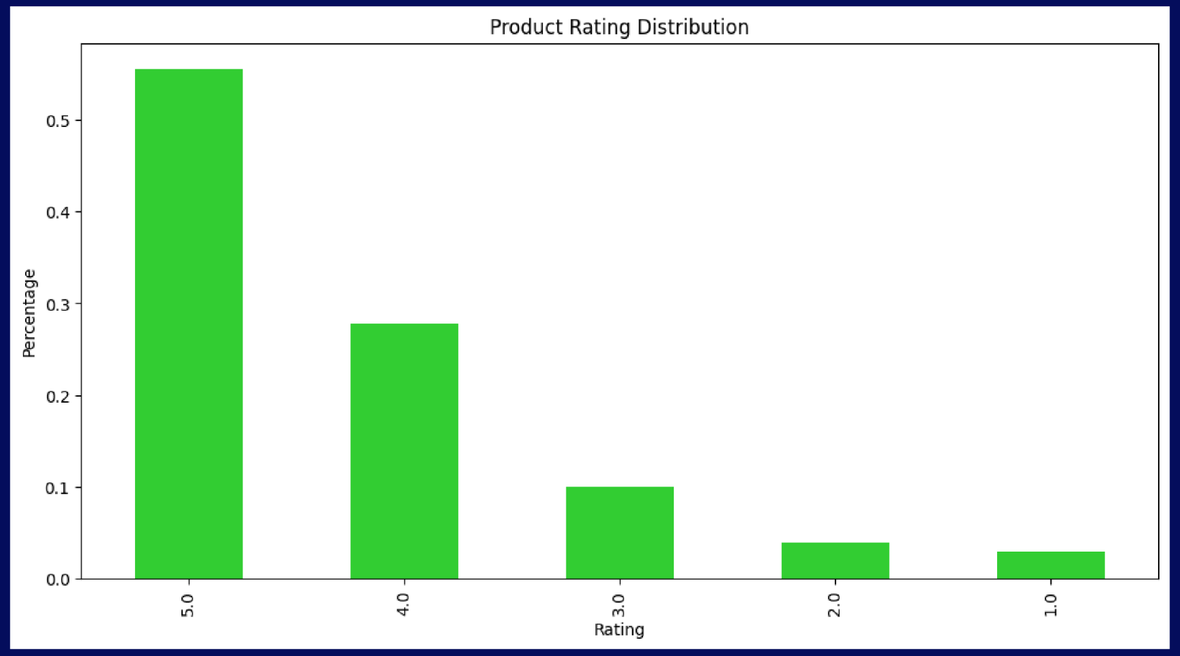
- The data contains 65,290 observations and 3 columns.
- Both the user_id and prod_id are object data types
- The rating of the products are numeric data types
- This was a clean dataset and shows no missing values in any row or column

```
count    65290.000000
mean         4.294808
std          0.988915
min          1.000000
25%          4.000000
50%          5.000000
75%          5.000000
max          5.000000
Name: rating, dtype: float64
```

- The average rating score is ~4.29 & maximum rating is 5
- The standard deviation of 0.988915 suggests some variability in the ratings; however, it's not extremely high, indicating that ratings are somewhat consistent
- The distribution indicates that a significant portion of products received the highest rating of 5.

EDA CONTINUED...

Rating Distribution - Bar Plot



- **High percentage of 5-star ratings indicates overall positive opinions.**
- **General satisfaction, but limited variability in ratings.**
- **Possible ceiling effect: Users hesitant to give lower ratings.**
- **Explore user reviews for a more comprehensive understanding.**

Unique Users & Items in dataset

- **65,290 observations for analysis.**
- **1,540 unique users suggest a diverse customer base.**
- **5,689 unique items imply a wide range of rated entities.**
- **Potential sparsity due to extensive items and users.**

RANK BASED MODEL

01

Calculate average rating and count of ratings for each product

02

Create a dataframe with calculated average and count of ratings and sort by average of ratings in descending order

03

Define a function to get the top n products based on highest average rating and minimum interactions

04

Recommend top 5 products with 50 minimum & 100 minimum interactions based on popularity

Top 5 with 50	['B001TH7GUU', 'B003ES5ZUU', 'B0019EHU8G', 'B006W8U2MU', 'B000QUUFRW']
Top 5 with 100	['B003ES5ZUU', 'B000N99BBC', 'B007WTAJTO', 'B002V88HFE', 'B004CLYEDC']

USER-USER SIMILARITY-BASED MODEL

```
RMSE: 1.0012
Precision:  0.855
Recall:  0.858
F_1 score:  0.856
```

```
RMSE: 0.9530
Precision:  0.847
Recall:  0.893
F_1 score:  0.869
```

DEFAULT PARAMETERS

HYPERPARAMETER TUNING

RMSE

- **Default**: **1.0012**
- **Tuned**: **0.9530**
- **Improvement**: Reduced RMSE in the tuned model indicates enhanced predictive accuracy.

RECALL

- **Default**: **0.858, 85%**
- **Tuned**: **0.893, 89%**
- **Improvement**: Increased recall in the tuned model signifies a better ability to capture relevant items.

PRECISION

- **Default**: **0.855, ~85%**
- **Tuned**: **0.847, ~84%**
- **Change**: Slight decrease in precision in the tuned model

F1 SCORE

- **Default**: **0.856, ~85%**
- **Tuned**: **0.869, ~86%**
- **Improvement**: Higher F1 score in the tuned model suggests a better balance between precision and recall.

USER-USER SIMILARITY-BASED MODEL

- The switch from **cosine similarity** in the default model to **MSD** during hyperparameter tuning resulted in improved performance:
- **RMSE**: Tuned model (MSD) achieved **lower RMSE,** indicating increased predictive accuracy.
- **Precision**: Both models maintained high precision, with a slight decrease in the tuned model.
- **Recall**: Tuned model showed **improved recall**, capturing a higher proportion of relevant items.
- **F1 Score**: The tuned model achieved a **higher F1 score**, indicating a better balance between precision and recall.
- Revised MSD model demonstrated improved accuracy (**lower RMSE**), enhanced ability to capture relevant items (**higher recall**), and a better overall balance between precision and recall (**higher F1 score**).

ITEM-ITEM SIMILARITY-BASED MODEL

DEFAULT PARAMETERS

- RMSE: 0.9950
- Precision: 83.8%
- Recall: 84.5%
- F1 Score: 0.841

```
RMSE: 0.9950
Precision:  0.838
Recall:  0.845
F_1 score:  0.841
```

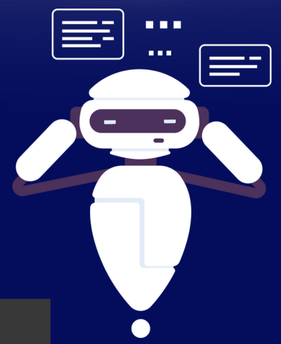
HYPERPARAMETER TUNING

- RMSE: 0.9575
- PRECISION: 84.7%
- RECALL: 89.3%
- F1 SCORE: 0.858

```
RMSE: 0.9530
Precision:  0.847
Recall:  0.893
F_1 score:  0.869
```

THE HYPERPARAMETER-TUNED MODEL (MSD) OUTPERFORMED THE DEFAULT MODEL WITH COSINE SIMILARITY, ACHIEVING ENHANCED PREDICTIVE ACCURACY (RMSE: 0.9575 VS. 0.9950). BOTH MODELS EXHIBITED THE SAME PRECISION (0.838), BUT THE TUNED MODEL DEMONSTRATED IMPROVED RECALL (0.88 VS. 0.845) AND A HIGHER F1 SCORE (0.858 VS. 0.841), INDICATING A BETTER BALANCE BETWEEN PRECISION AND RECALL.

MATRIX FACTORIZATION BASED MODEL



DEFAULT PARAMETERS

- RMSE: 0.8882
- Precision: 85.3%
- Recall: 88%
- F1 Score: 0.866

```
RMSE: 0.8882
Precision:  0.853
Recall:  0.88
F_1 score:  0.866
```

HYPERPARAMETER TUNING

- RMSE: 0.8811
- PRECISION: 85.3%
- RECALL: 88.2%
- F1 SCORE: 0.867

```
RMSE: 0.8811
Precision:  0.853
Recall:  0.882
F_1 score:  0.867
```

THE HYPERPARAMETER-TUNED MODEL RMSE SCORE (0.8811) IMPROVED SLIGHTLY OVER THE DEFAULT (0.8882), SUGGESTING ENHANCED PREDICTIVE ACCURACY. PRECISION REMAINED THE SAME AT 85.3%, WHILE RECALL INCREASED FROM 88% TO 88.2%, AND THE F1 SCORE IMPROVED FROM 0.866 TO 0.867. OVERALL, THE TUNED SVD MATRIX FACTORIZATION-BASED MODEL ACHIEVES ACCURATE PREDICTIONS WITH A REFINED BALANCE BETWEEN PRECISION AND RECALL.

CONCLUSIONS

- The optimal approach for predictive accuracy in product recommendations—whether **item-item collaborative filtering**, **user-user collaborative filtering, item-item collaborative filtering,** or **matrix factorization (SVD)**—depends on factors like data characteristics and system goals.

- The result for SVD is only better than the user-user similarity based recommendation system and the not the optimized one.
- Building the item-item similarity-based recommendation system proved to have the highest predictive accuracy. Predicting a products rating closest to how the user rated it. 4.67/5
- The evaluation metrics suggest that item-item model has the highest performance when used to predict recommended products for users.

RECOMMENDATIONS

- Some approaches to enhance the performance of item-item similarity model include:
  - Explore additional features or attirbutes that can contribute to a more nuanced understanding of item similarities. Perhaps metadata, tags, or other relevant information
  - Explore hybrid approaches by combining item-item collaborative filtering with content-based methods
  - The difference in prediction accuracy for the same product can be attributed to the individual user preferences, interaction history, and the inherent sparsity of the data.
  - Clustering techniques to improve the accuracy of the model
  - Monitoring the consistency in the predicted ratings across different products for a specific user suggests that the model is assigning similar ratings based on the user's preferences or historical interactions

THANK YOU!