Recommendation Systems

Amazon User Recommendation Solution – MIT NCAIML

July 2, 2023 - Edward Lang

Recommendations Based
on Previous User Ratings:

Electronic Products

Disclaimers & Notices

The following case study was created for educational purposes only. All data used has been anonymized.

Sources utilized in preparing this presentation:

https://www.amazon.science/blog/using-graph-neural-networks-to-recommend-related-products
https://financesonline.com/amazon-statistics/
https://www.zippia.com/advice/amazon-statistics/
https://www.mageplaza.com/blog/product-recomendation-how-amazon-succeeds-with-it.html
https://boostcommerce.net/blogs/all/lesson-learned-from-amazon-product-recommendation
https://recostream.com/blog/amazon-recommendation-system

Agenda

Executive Summary

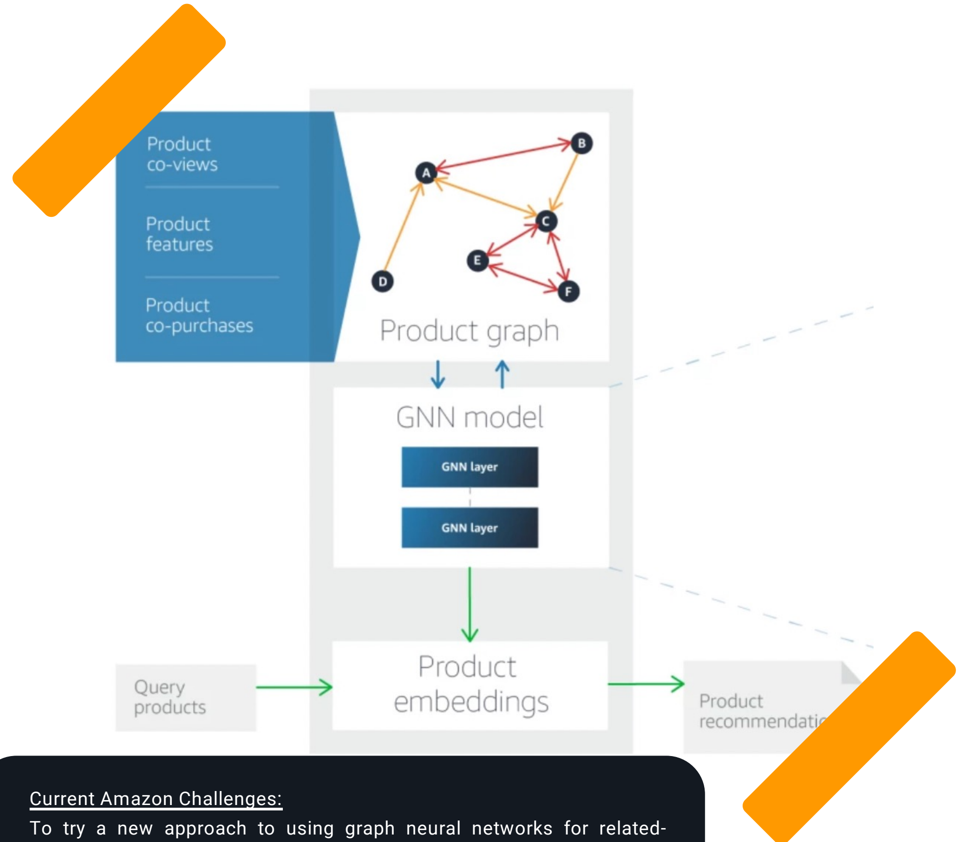
Business Problem: To try and improve Amazon's recommendation systems by building and comparing both an Item-Item and User-User system.

Data Insights: We have over 1 million records with ratings between 1-5 on electronic products.

Methodology: utilize collaborative filtering techniques and test various models.

Recommended Solution: Based on the initial parameters provided, we recommend the User-User system.

Further Considerations: With the addition of the user's comments, we could have also looked at a content-based model for comparative purposes.



Current Amazon Challenges:

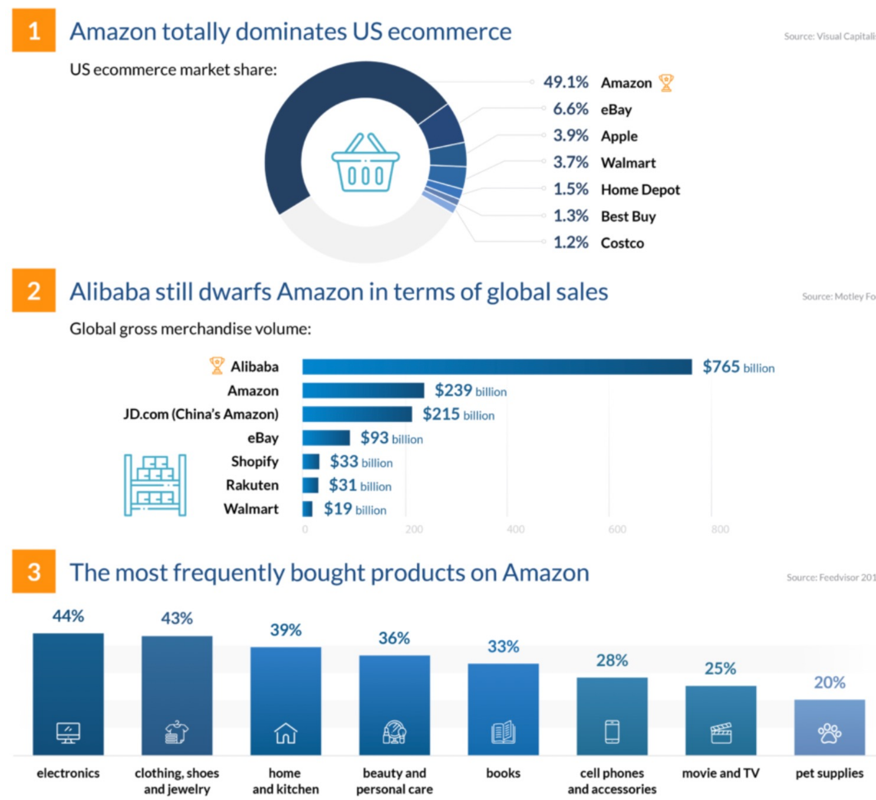
To try a new approach to using graph neural networks for related-product recommendation produces two embeddings of every graph node: one that characterizes its role as the source of a recommendation and one that characterizes its role as the target.

Business Problem: Overview

There is no denying that Amazon is an eCommerce GIANT. 2nd to Alibaba, it is the market leader in the US, including electronics.

Some quick stats:

- Amazon's share of all e-commerce sales in the U.S. hit a whopping 56.7% in 2021. (Zippia, July 2023)
- Amazon has a catalog of 12 million products across all categories and services. (RepricerExpress, 2021)
- The Amazon product category with the most number of global keywords is electronics (70 keywords) followed by home (12), media (7), miscellaneous (6), clothing (4), and food (1). (Visual Capitalist, 2020)
- In the U.S., electronics was the most purchased product category on Amazon Prime Day at 32%, followed by household essentials (22%), health and beauty (21%), and toys and video games (21%). (Numerator, 2020)
- 63.5% of Amazon traffic comes from the U.S. (Amazon, 2021)



Source: https://financesonline.com/amazon-statistics/

Business Problem: Details

History

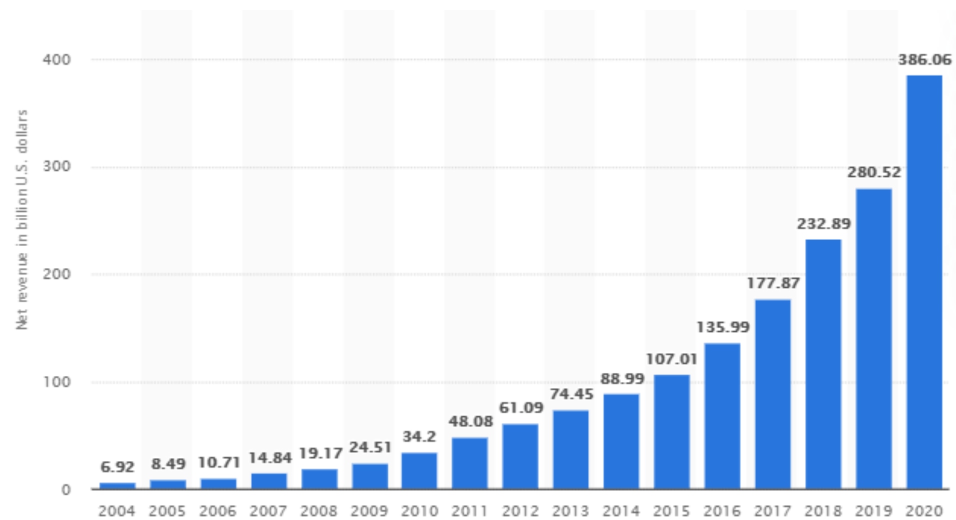
Amazon is no stranger to collaborative filtering for a recommendation system. In fact, they pioneered various methods of collaborative filtering, including a published article Recommendations: Item-to-Item Collaborative Filtering in 2003, which later won an award from the Institute of Electrical and Electronics Engineers (IEEE) as the published article that best stood the test of time.

In contrast to content-based filtering, Amazon discovered that other filtering techniques, such as the collaborative filtering done in this study, produces better results, specifically:

- **diversity** - group filtering generates a more diverse list of recommended products, offering customers a wider choice,
- **randomness** - recommendations are much more likely to positively surprise customers and show them a product of interest, which they may have never considered, and
- **randomness** - these methods can more effectively present to customers offers they would most likely be interested in.

(Source: https://recostream.com/blog/amazon-recommendation-system)

According to a McKinsey study, **up to 35%** of Amazon's sales are generated thanks to the proprietary product **recommendation algorithm.**



Source: https://www.statista.com/statistics/266282/annual-net-revenue-of-amazoncom/

Problem Definition

As the US eCommerce dominate consumer site, it is critical, with so much choice in the Amazon market, to recommend useful products to consumers.

Background

Data Collected – Amazon uses 2 sources

- general data related to products and users; and
- data about the relationships and dependencies between them.

Amazon's Recommendation algorithm 3 main types of analysis

- User-Product;
- Product-Product; and
- User-User

Addition data utilized in the current Amazon rrecommendation aalgorithm

- User behavior;
- User demographics; and
- Product attributes.

Goals

Initial Objectives

- Gain insights into the relationships between users and product ratings; and
- Build a Recommendation System model using different techniques (I-I & U-U) to try and predict both ratings and recommend electronics to other users
- Compare different models and select the best performing (add'l analysis is in the Appendix).

Not considered in this study

- Since none of the additional data utilized by Amazon is provided, these recommendation models are focused only only the product ratings provided by users

Problem Definition

Business Problem – how to recommend relevant products to users based on either the ratings of similar type users or based on the consumers purchase history and make recommendations of similar items.

Amazon's Top selling electronic products by *Ranking (the simplest form of recommendation).*

Solution Design

The following methodology is being utilized for this study.

Exploratory Data

Initial examination of the data to try and understand if there are any hidden patterns in the data set, distributions, and/or any relationship between the variables

Data Preparation

The initial pre-processing needed on the data for analysis, i.e. filtering the records, aggregation, and setting rating minimums for model building.

Recommendation Methods

Build different Recommendation Methods to recommend electronics to users

Hyper-parameter tuning (wherever needed)

Build the best model based on the optimum parameters

Conclusion

To understand the difference between and be able to recommend the best performing recommendation algorithms

Solution Design: Methods Applied

Models Performed Beyond Requested Study (Results in Appendix)

Assignment Requirements

Rank-based Recommendation System: This is the simplest method of creating recommendation systems, where we assume that all customers have similar preferences.

User-User Collaborative Filtering: an item is recommended to a user based on user-user similarity, by looking at the items used by similar users.

Item-Item Collaborative Filtering: an item is recommended to the user simply based on item-item similarity with items this user has already used.

Singular Value Decomposition (SVD) Based Collaborative Filtering:, Matrix Factorization is applied to fill user-item interaction matrix in order to make recommendations.

Performance Metrics

RMSE: RMSE or Root Mean Squared Error, measures the average difference between a statistical model's predicted values and the actual values. So for our purposes, that would be predicted ratings. We have been asked to consider RMSE first as it is more sensitive to outliers. Hence, the lower the RMSE, the better the model.

MAE: MAE stands for Mean Absolute Error and is a measure of errors between paired observations. For our study, this is the average absolute difference between the predicted rating and the actual rating. We will consider this result second – lower is better.

NMAE: NMAE or Normalized Mean Absolute Error is used to facilitate the comparison regarding MAE of datasets with different scales. This is a normalized version of MAE. f NMAE, the better the model. We consider 3rd – lower is better.

Exploratory Data Analysis: Data Definition

There are **three features** present in this dataset:

I. User_Id – this randomized attribute identifies each unique Amazon customer that rated an electronic product
II. Product_Id - this randomized attribute identifies each unique Amazon electronic product that was rated
III. Rating – this attribute is the actual rating, between 1-5, that each user gave to a specific electronic product

Dataset Features

User Id	Every user identified with a unique id
Product Id	Every product identified with a unique id
Rating	The rating of the product by the user

EDA: Data Exploration - Data Overview

- **1,048575 Rows**: Each row in the dataset represents a rating of an electronic product.
- **3 Columns**: The columns/attributes in the dataset contain the necessary details about the User ID, Product ID and Rating

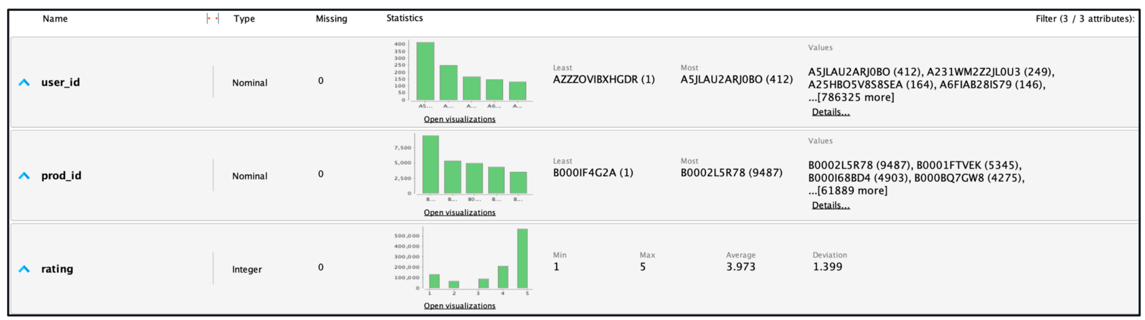
ow No.	user_id	prod_id	rating
1	A2CX7LUOHB2NDG	321732944	5
2	A2NWSAGRHCP8N5	439886341	1
3	A2WNBOD3WNDNKT	439886341	3
4	A1GI0U4ZRJA8WN	439886341	1
5	A1QGNMC6O1VW39	511189877	5
6	A3J3BRHTDRFJ2G	511189877	2
7	A2TY0BTJOTENPG	511189877	5
8	A34ATBPOK6HCHY	511189877	5
9	A89DO69P0XZ27	511189877	5
10	AZYNQZ94U6VDB	511189877	5
11	A1DA3W4GTFXP6O	528881469	5
12	A29LPQQDG7LD5J	528881469	1
13	AO94DHGC771SJ	528881469	5
[illegible]	[illegible]	528881469	1
[illegible]	[illegible]	528881469	4

Data set
examples

EDA: Data Exploration - Summary Statistics

- There are **no missing** attributes.
- The ratings system is between **1 and 5**.
- The **average of all the ratings** given by users is **3.973**. This indicates that the electronics being reviewed on Amazon are generally getting a good ratings from consumers.

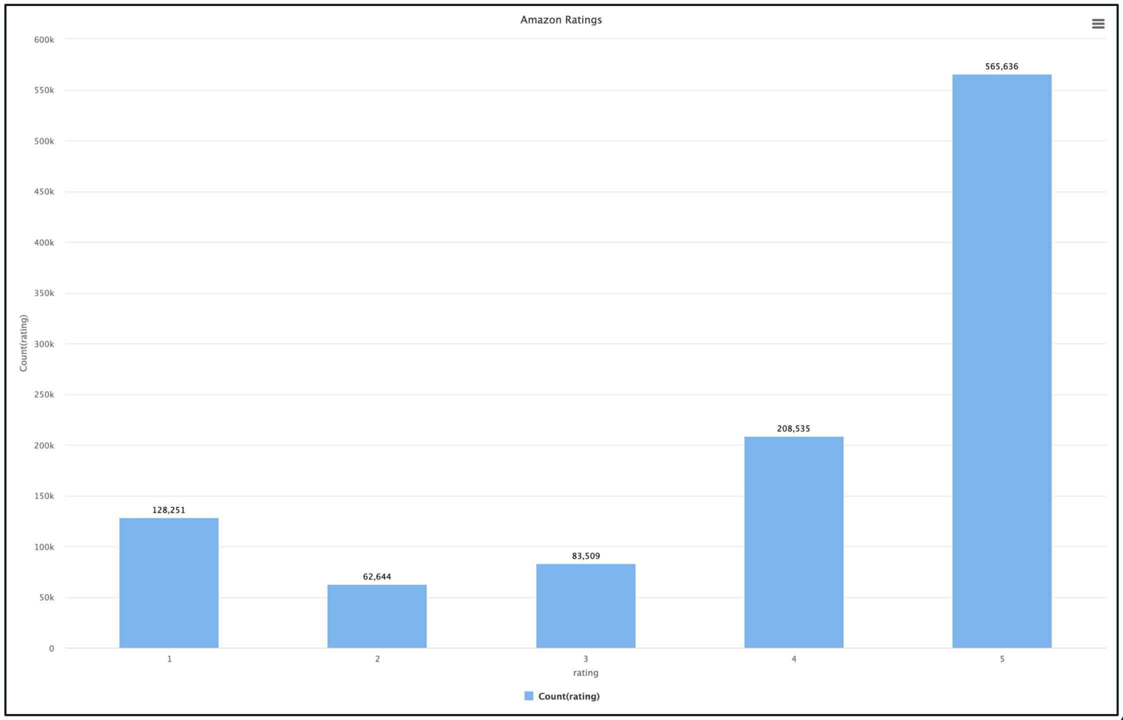
Data set basic statistics – Rapid Miner



EDA: Data Exploration - Ratings Distribution

- Plotting the ratings reveals that ~55% of Amazon customers have rated electronics purchased a 5-star rating and nearly 75% give it a 4-star rating or higher. This indicates most customers are very satisfied with their purchases

- ~7% of customers however (a small proportion) have rated the electronics low, i.e., a 1 or 2.



EDA: Data Exploration - Ratings Count by Product

- This table shows those electronics with the **highest counts** of ratings in the dataset.
- There are a total of **61,893 unique products** of electronics available in the dataset.
- Product **"B0002L5R78"** received the highest number of ratings of any other products (almost 2x1).

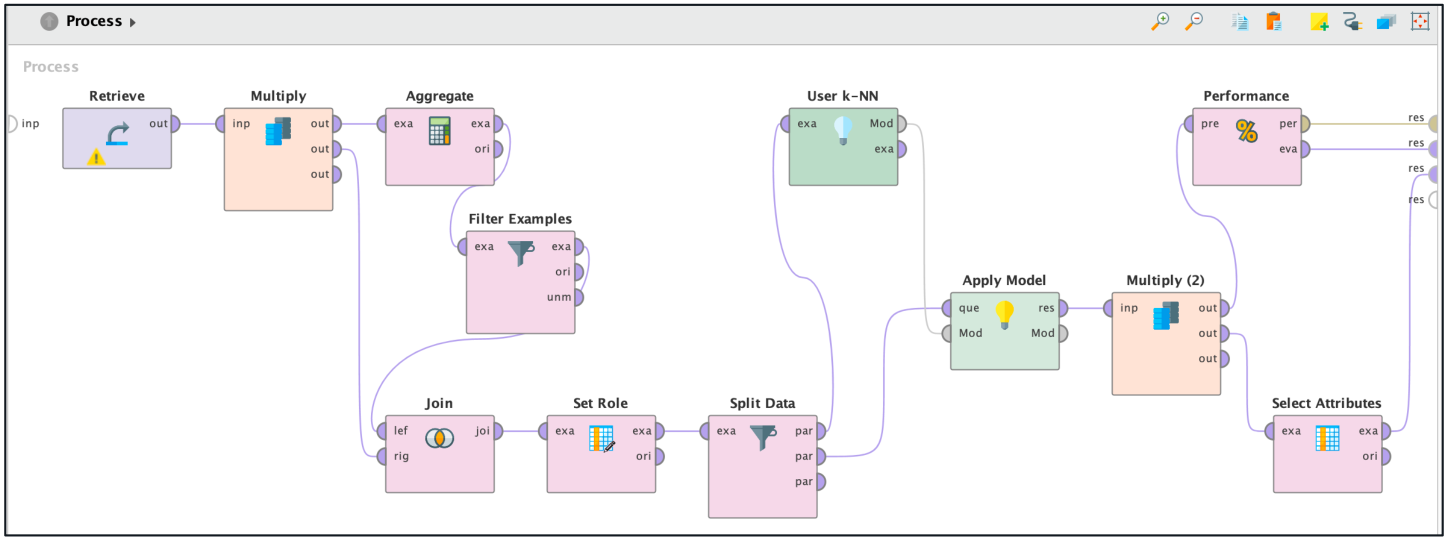
Row No.	prod_id	count(prod_id) ↓
30276	B0002L5R78	9487
24439	B0001FTVEK	5345
61285	B000I68BD4	4903
46504	B000BQ7GW8	4275
14183	B00007E7JU	3523
45867	B000BKJZ9Q	3219
45092	B000B9RI14	2996
43023	B000A6PPOK	2828
14780	B00007M1TZ	2608
5130	B00004ZCJE	2547

EDA: Data Exploration - Count by Users

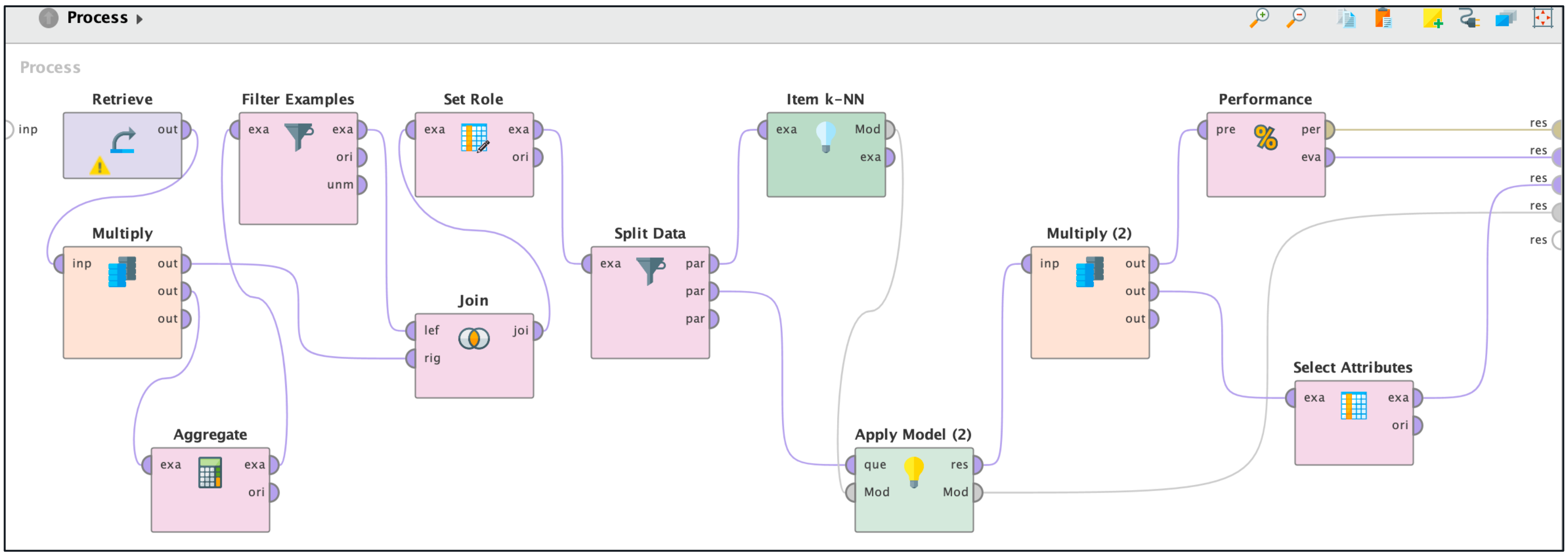
Row No.	user_id	count(user...
1	A5JLAU2ARJ0BO	412
2	A231WM2Z2JL0U3	249
3	A25HBO5V8S8SEA	164
4	A6FIAB28IS79	146
5	AT6CZDCP4TRGA	128
6	AKT8TGIT6VVZ5	122
7	A11D1KHM7DVOQK	112
8	A2B7BUH8834Y6M	103
9	A3OXHLG6DIBRW8	95
10	A203OCQQ12MAVT	90

- This table is a quick check to show the User IDs for users who have provided the highest number of reviews for Amazon electronics in the dataset.
- There are a total of **786,329 unique users**, 412 is the highest number of reviews provided by a single user.
- As per the number of unique users and businesses, there is a possibility of 786,329 x 61,893 = **48,668,260,797** ratings. However, we only have **1,045,575** ratings (around **0.0021%**), a very small amount of the total possible ratings.

User-User Similarity-Based Model



Item-Item Similarity-Based Model



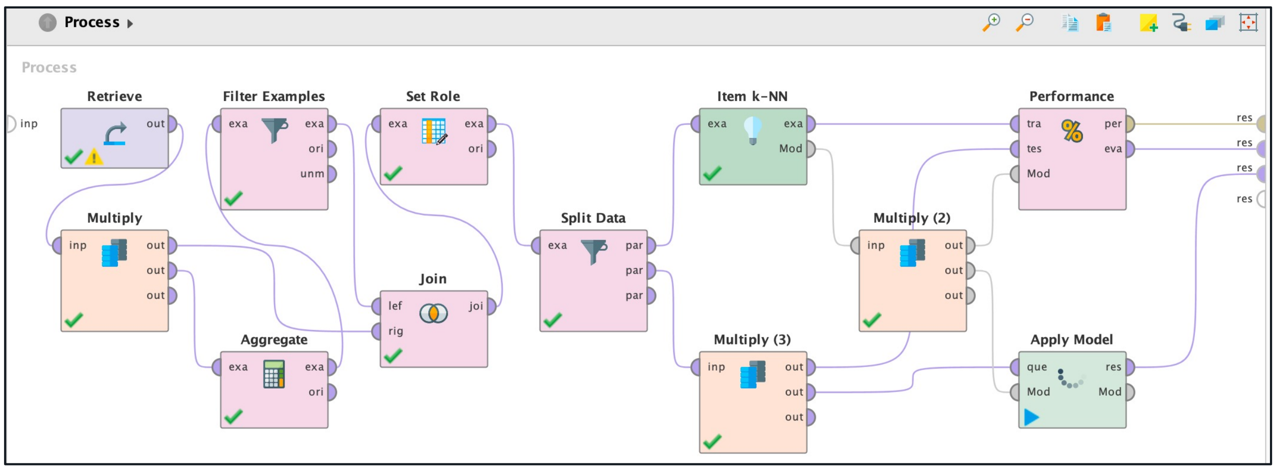
Model Comparisons

Model Performance Metrics	RMSE	MAE	NMAE	Runtime
Item-Item Collaborative	1.271	0.981	0.245	5 secs
User-User Collaborative	1.155	0.892	0.223	5 secs

Here are the performance metrics of the 2 different collaborative filtering-based methods asked to be performed using the Amazon dataset:

In our tests – after running the ***User-User Similarity Based Recommendation System*** and the ***Item-Item Similarity-Based Recommendation System*** with users who have rated over 100 products, the User-KNN model with k=90, and minimum rating of 1 and range of 4, using cosine similarity as the similarity metric; ***the User-to-User Process*** produced the best results.

Item Recommendation



Performance Vector (Performance)
Result not stored in repository.

```
PerformanceVector:
AUC: 0.395
prec@5: 0.001
prec@10: 0.001
prec@15: 0.001
NDCG: 0.110
MAP: 0.005
```

Items Recommended: Optimum Model

- As requested, here is the output for the top 5 recommended electronic products based on our item recommendation model.
- As previously mentioned, with the data set, the user IDs and product IDs have been anonymized.

user_id	item_id	rank
1309	60505	1
1309	6691	2
1309	2548	3
1309	16436	4
1309	41718	5
62644	61838	1
62644	2692	2
62644	8047	3
[illegible]	[illegible]0953	4
[illegible]	[illegible]119	5

Top 5 Recommendations

Examples for 2 users

Conclusions

With our initial analysis, we found that most ratings fell between **4 and 5 (75%)** and the overall average rating from the data set was **3.973.**

- Based on our findings, Amazon can continue to aid its customers by continuing to utilize the **User-User recommendation model** as shown in this study (Rapid Miner files available upon request).

- It has been observed in our initial study that the User-User Collaborative Model produced the best performance based on RMSE, MAE, and NMAE. Even after testing more parameters and additional techniques, including training with more or less records, **the User-User model performs the best**.

- Amazon should continue using Item-Item Collaborative Filtering models to help make personalized electronic recommendations to enhance **discovery**, increase **loyalty** and **engagement** from customers.

Other Considerations:

- If we were given the actual user reviews, we could have considered a Content-Based Recommendation System

- If we were given other products, we could have seen the relationship between product categories on a user-by-user basis.

Appendix

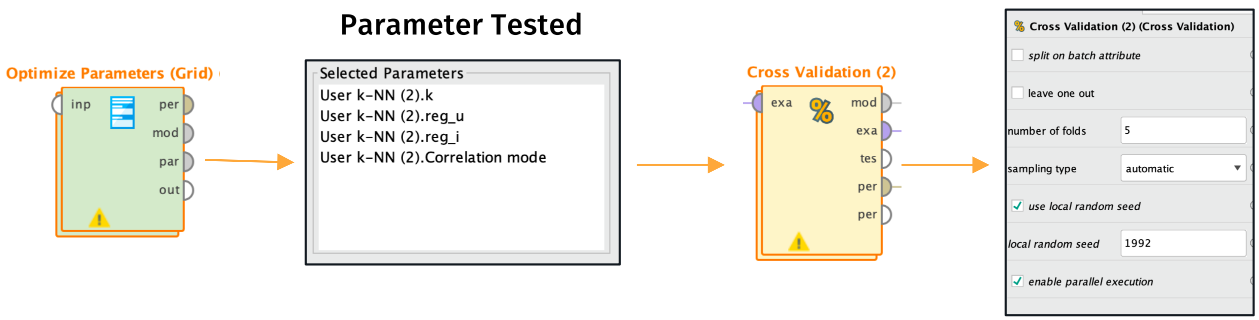
Model Results: Rank-Based Recommendation

Row No.	prod_id	count(rating) ↓	average(rating)
30276	B0002L5R78	9487	4.449
24439	B0001FTVEK	5345	4.007
61285	B000I68BD4	4903	3.502
46504	B000BQ7GW8	4275	4.553
14183	B00007E7JU		
45867	B000BKJZ9C		
45092	B000B9RI14		
43023	B000A6PPO		
14780	B00007M1T		
5130	B00004ZCJE		

Row No.	prod_id	count(rating)	average(rat... ↓
1	059400232X	3	5
3	094339676X	1	5
5	1039869017	2	5
8	1182702627	3	5
31	1575839415	1	5
33	1593152523	2	5
34	1594243034	1	5
37	1604550945	1	5
38	1610130804	1	5
49	1615598790	1	5

- This model is the most basic form of a recommendation system by simply recommending the highest-rated product within the category to a new visitor.
- This type of recommendation model is a non-personalized technique.
- This Rank-based model was based on users who had at least 100 rated items are taken into consideration.
    - We can see that the product with the most ratings had an average of 4.449.
    - Also, notice the ten items that average a rating of 5, despite having a small number of ratings overall.

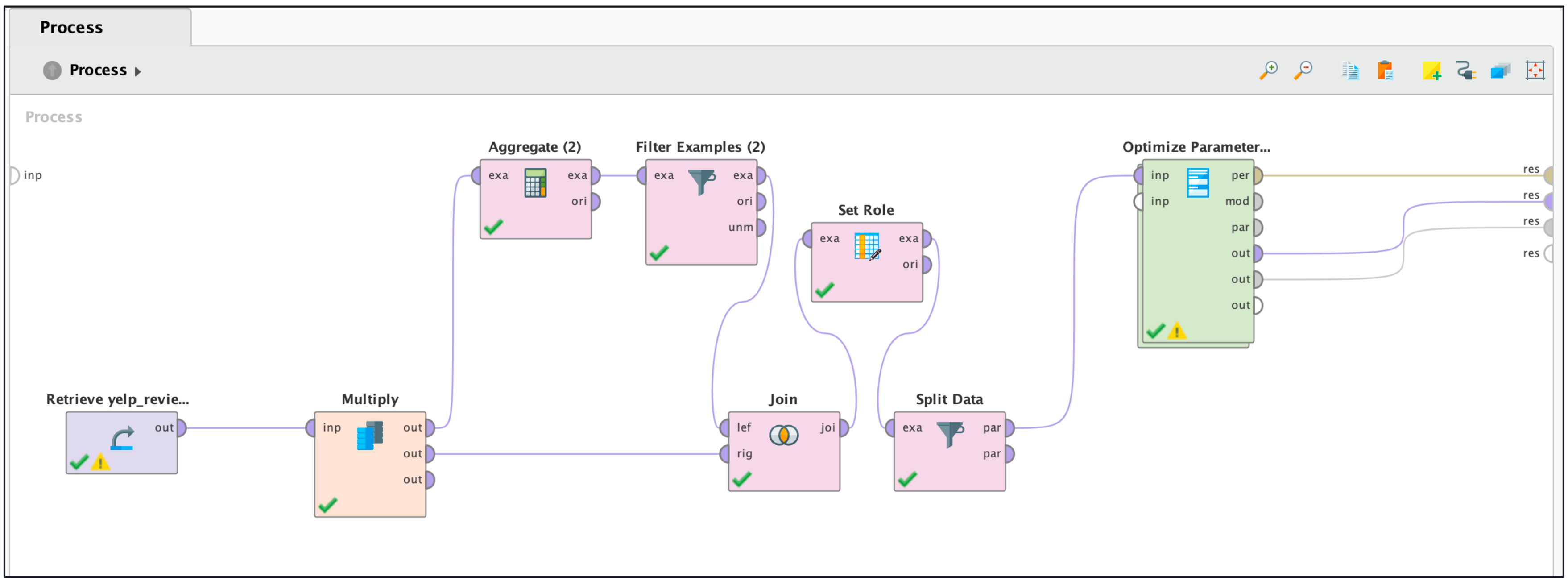
Model: Collaborative Filtering Using the Optimize Parameters Operator for Hyperparameter Tuning



Pearson and Cosine similarity are both measures of similarity between two sets of data but used in different ways. Pearson is a measure of linear correlation between two sets of data. It is commonly used to measure the correlation between two continuous variables, such as the relationship between height and weight. Cosine is a measure of similarity between two non-zero vectors. It is commonly used in NLP such as measuring the similarity between two documents. So, Pearson is best used for continuous data, while Cosine similarity is best used for discrete data.

(Source: https://www.quora.com/In-what-scenario-is-using-Pearson-correlation-better-than-Cosine-similarity)

Model: Collaborative Filtering
User-User Hyperparameter Tuning



Model: Collaborative Filtering User-User Hyperparameter Tuning RESULTS

Attribute	Type	Missing	Statistics		
Label rating	Integer	0	Min 1	Max 5	Average 4.090
User identification user_id	Polynominal	0	Least AZZZOVIBXHGDR (0)	Most A5JLAU2ARJ0BO (285)	Values A5JLAU2ARJ0BO (285), A231WM2Z2JL0U3 (170), ...[786327 more]
Item identification prod_id	Polynominal	0	Least B000IF51UQ (0)	Most B00004S9AK (2)	Values B00004S9AK (2), B00004SCKA (2), ...[61891 more]
count(rating)	Integer	0	Min 103	Max 412	Average 231.565
average(rating)	Real	0	Min 3.125	Max 4.963	Average 4.093

Optimized Parameter GRID Stats

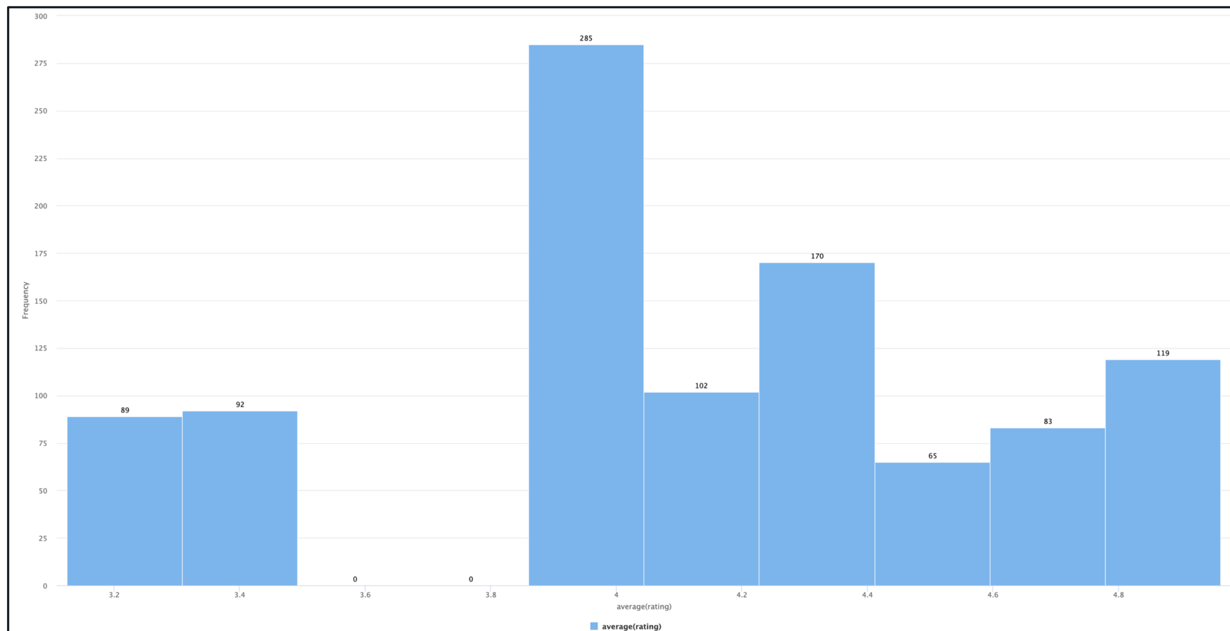
Fewer Records

Row No.	rating	user_id	prod_id	count(rating)	average(rat...
1	2	A11D1KHM...	B000001OMN	112	3.125
2	1	A11D1KHM...	B00000J1G6	112	3.125
3	4	A11D1KHM...	B00000J9Z7	112	3.125
4	1	A11D1KHM...	B00000JBYW	112	3.125
5	4	A11D1KHM...	B00000JCTD	112	3.125
6	4	A11D1KHM...	B00001P3XM	112	3.125
7	4	A11D1KHM...	B00001W0D4	112	3.125
8	3	A11D1KHM...	B00001ZWRV	112	3.125
9	1	A11D1KHM...	B00004TZK6	112	3.125
10	3	A11D1KHM...	B00004Z0BN	112	3.125

ExampleSet (1,005 examples, 3 special attributes, 2 regular attributes)

Model: Collaborative Filtering User-User Hyperparameter Tuning RESULTS

- Observation that the average rating range was reduced to those users with an average rating between 3.1 and 4.9 as the following table shows the distribution.

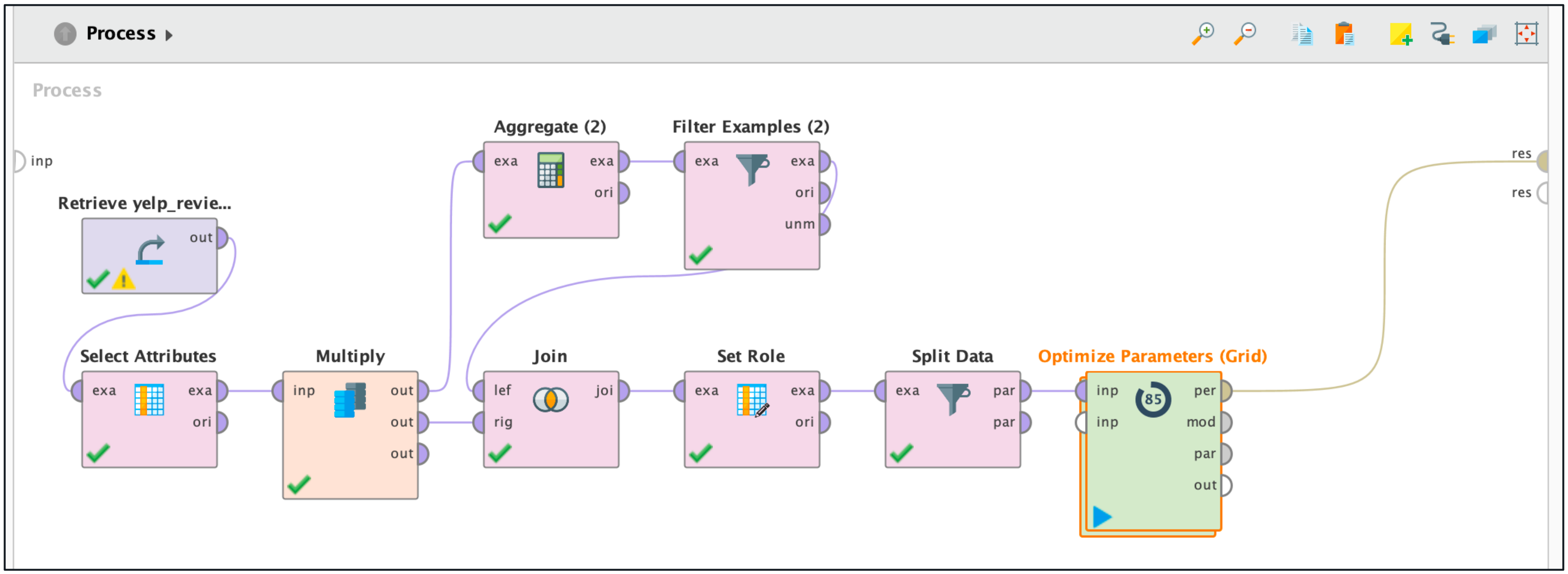


Model: Collaborative Filtering User-User Hyperparameter Tuning RESULTS

iteration	User k−NN (2).k	User k−NN (2).reg_u	User k−NN (2).reg_i	User k−NN (2).Correlation mode	RMSE
351	40	1	6	pearson	1.132
1	40	1	1	pearson	1.173
176	40	6	3	pearson	1.133
701	40	1	1	cosine	1.173
526	40	6	8	pearson	1.133

Optimized Parameter GRID Output

Model: Collaborative Filtering
Item-Item Hyperparameter Tuning



Model: Collaborative Filtering Item-Item Hyperparameter tuning RESULTS

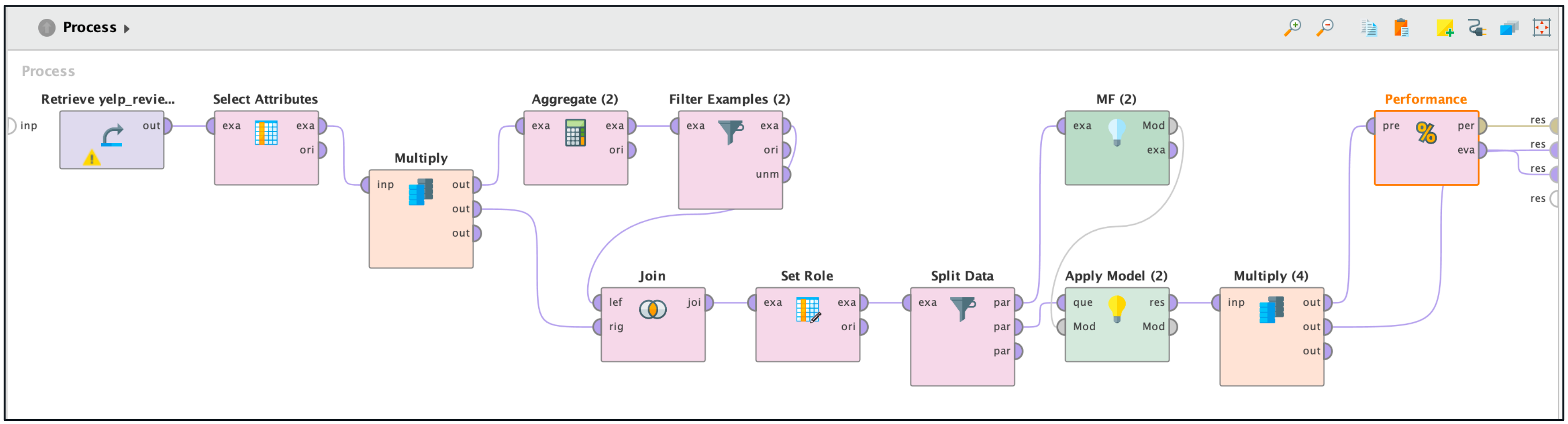
Row No.	RMSE	MAE	NMAE
1	0.988	0.752	0.188
2	1.149	0.915	0.229
3	1.130	0.844	0.211
4	1.285	0.985	0.246
5	1.137	0.862	0.216

By Rating

Optimized Parameter GRID Output

iteration	Item k-NN (2).k	Item k-NN (2).reg_u	Item k-NN (2).reg_i	RMSE ↑
37	50	1	5	1.138
38	60	1	5	1.138
39	70	1	5	1.138
40	80	1	5	1.138
41	90	1	5	1.138

Model: Collaborative Filtering
SVD Similarity-Based

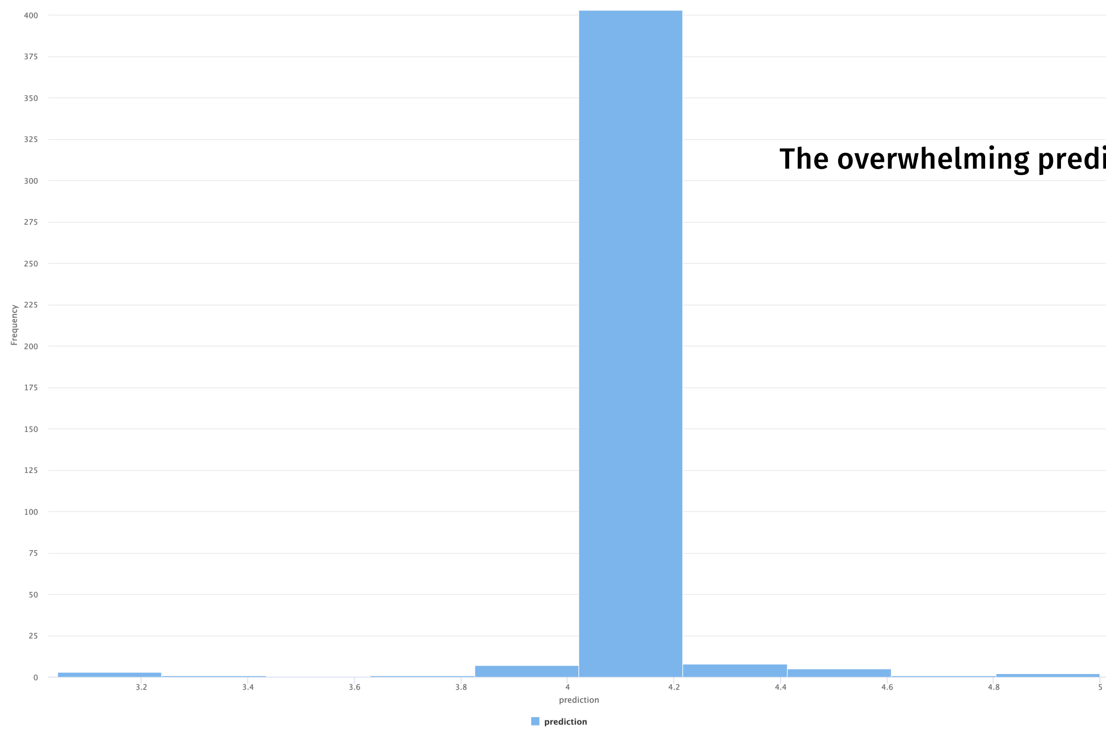


Model: Collaborative Filtering SVD Similarity-Based RESULTS

Row No.	rating	user_id	prod_id	count(rating)	average(rating)	predicti...
201	5	A5JLAU2ARJ...	B000066E70	412	3.871	5
199	5	A5JLAU2ARJ...	B0000658CG	412	3.871	4.814
274	2	A5JLAU2ARJ...	B00026IN1U	412	3.871	4.645
347	5	A6FIAB28IS79	B000BKJZ9Q	146	4.137	4.586
189	1	A5JLAU2ARJ...	B00004YKDQ	412	3.871	4.494
191	5	A5JLAU2ARJ...	B00005LEN4	412	3.871	4.477
98	3	A231WM2Z...	B0000APSKB	249	4.309	4.467
106	5	A25HBO5V...	B00005ICE1	164	4.963	4.426
275	4	A5JLAU2ARJ...	B0002B80EA	412	3.871	4.337
161	3	A2B7BUH88...	B0000899ZA	103	4.417	4.326

SVD Model Output

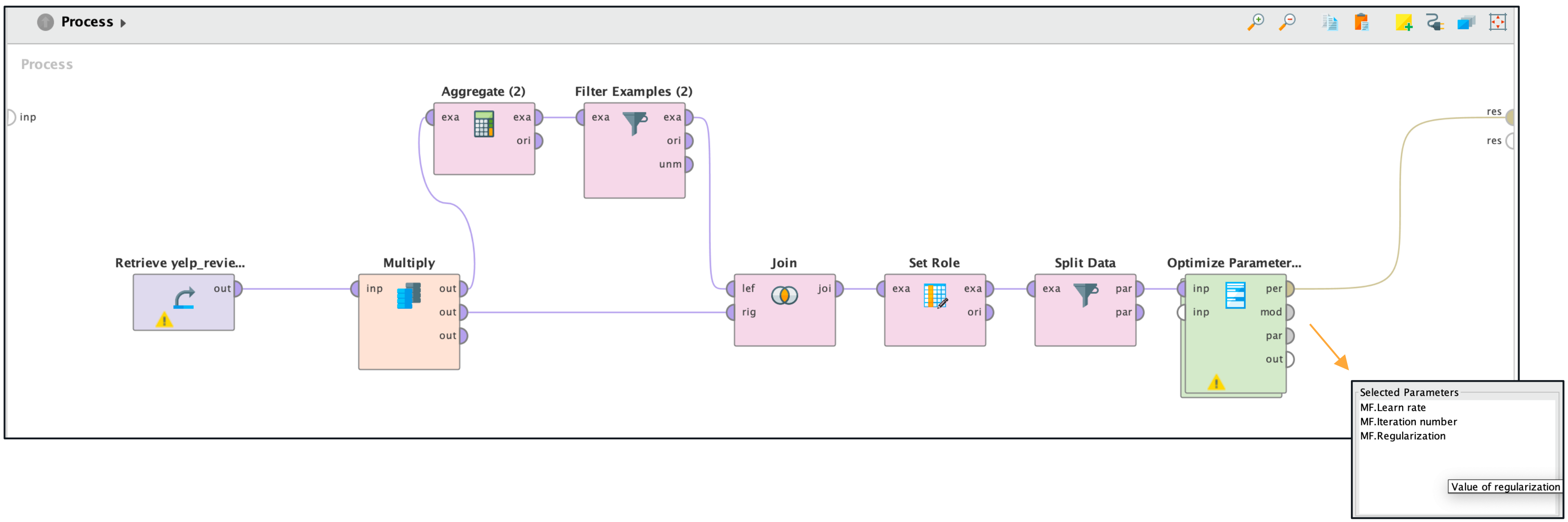
Model: Collaborative Filtering SVD Similarity-Based RESULTS



The overwhelming predicted rating was ~4.1

Model: Collaborative Filtering

SVD Hyperparameter Tuned

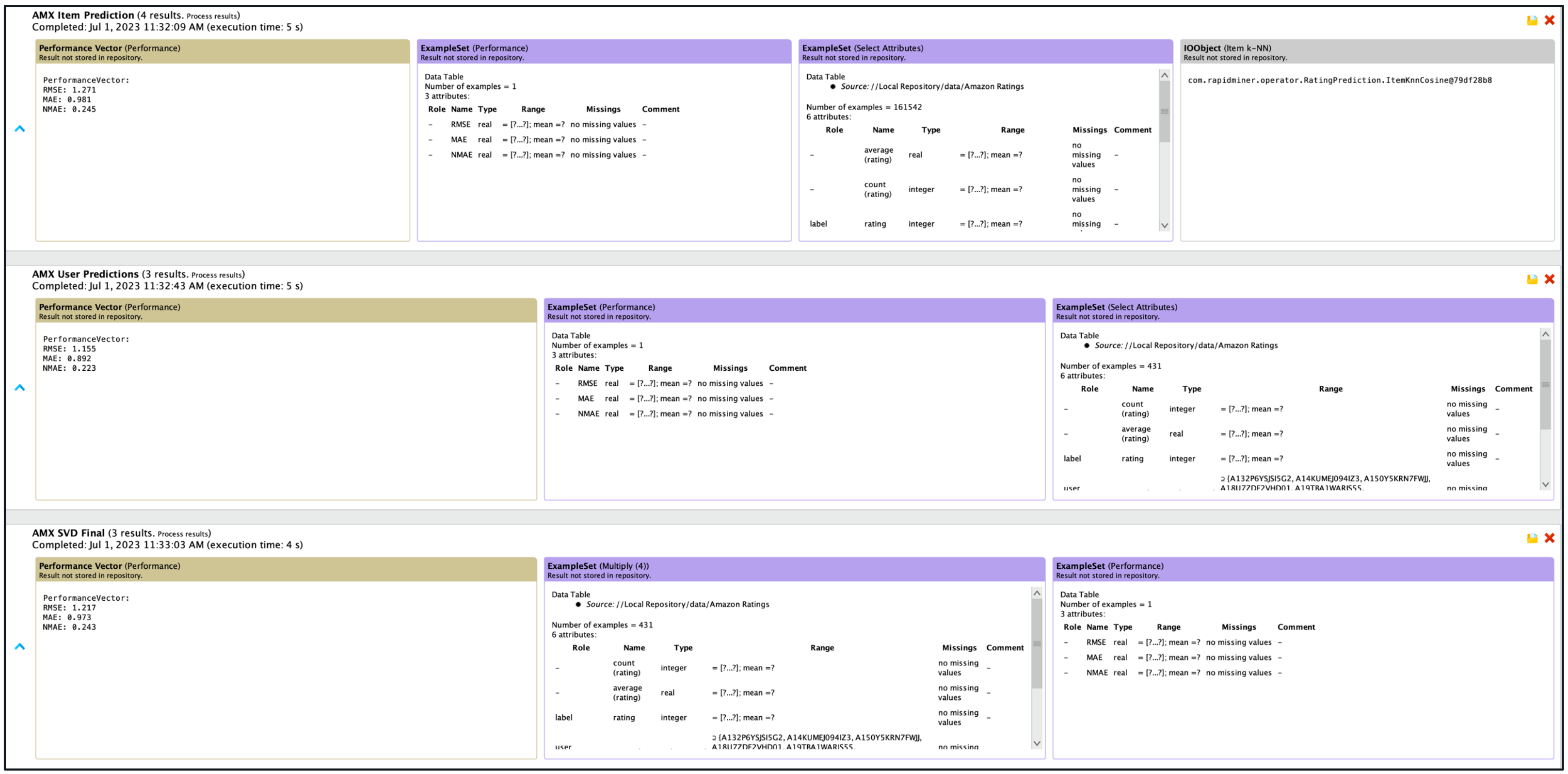


Model: Collaborative Filtering SVD Hyperparameter Tuned RESULTS

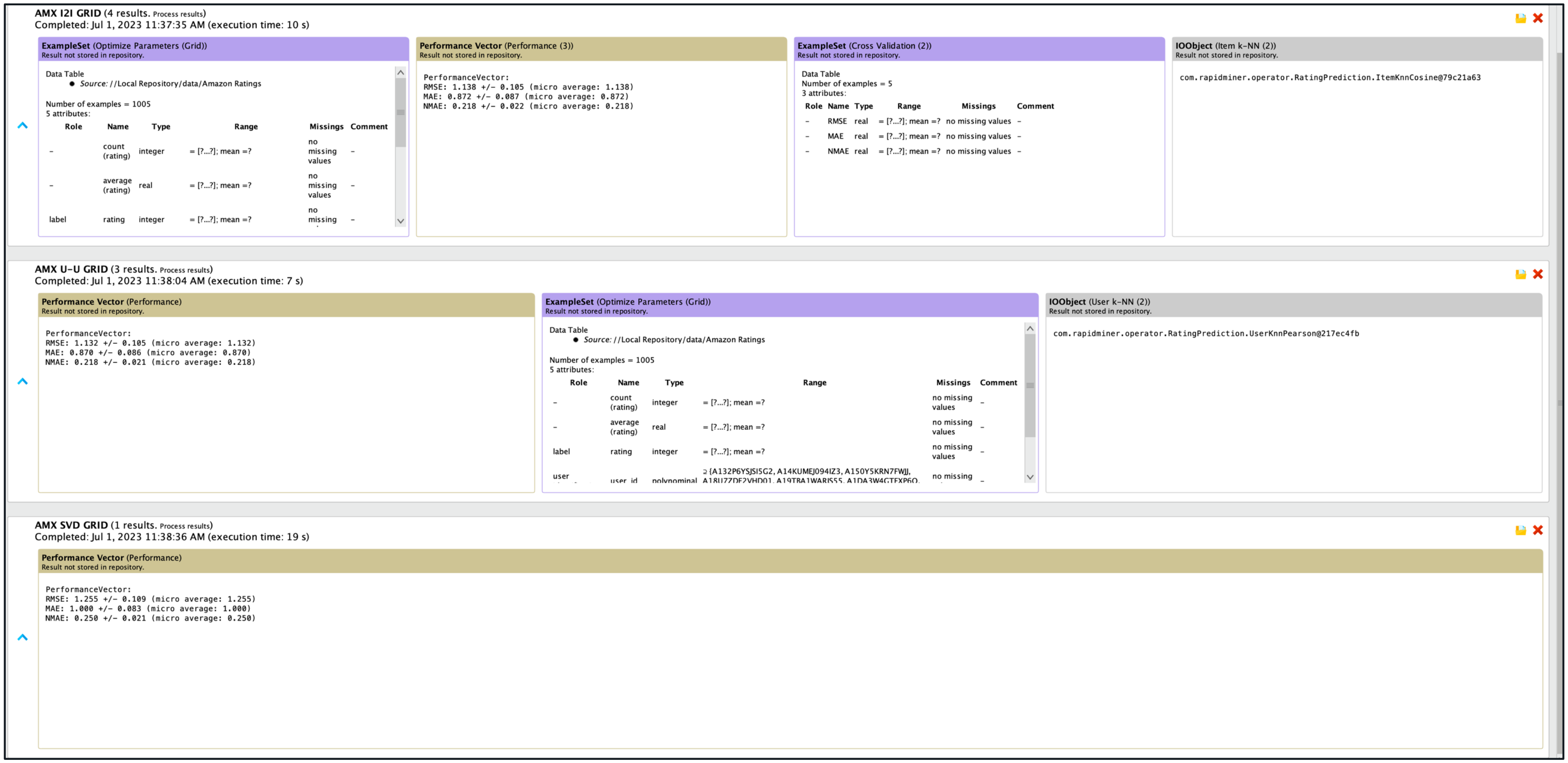
iteration	MF.Learn rate	MF.Iteration number	MF.Regularization	RMSE ↑
71	0.010	18	0.006	1.257
347	0.070	42	0.026	1.258
84	0.040	26	0.006	1.258
243	0.030	10	0.021	1.258
340	0.100	34	0.026	1.259

Optimized Parameter GRID Output

Models: I-I, U-U & SVD RESULTS



Models: I-I, U-U & SVD Optimization GRID RESULTS



Model Results: Collaborative Filtering

Model Name/Performance metrics	RMSE	MAE	NMAE
Item-Item Collaborative	1.271	0.981	0.245
Item-Item Collaborative (Tuned)	1.138	0.872	0.218
User-User Collaborative	1.155	0.892	0.223
User-User Collaborative (Tuned)	**1.132**	**0.870**	**0.218**
SVD	1.208	0.966	0.242
SVD (Tuned)	1.257	0.998	0.250

Here are the performance metrics of different collaborative filtering-based methods applied to the Amazon dataset:

In our tests – after running the ***User-User Similarity Based Recommendation System*** and the ***Item-Item similarity-based recommendation system*** with users who have rated over 100 products, the User-KNN model with k=90, and minimum rating of 1 and range of 4, using cosine similarity as the similarity metric; ***the User-to-User Optimization Process*** produced the best results with the following settings:

- User k-NN (2).k: 40
- User k-NN (2).reg_u: 1
- User k-NN (2).reg_i: 6
- User k-NN (2).Correlation mode: pearson