

Amazon Recommendation System

Jin-hoon Chung
September 30, 2022

The Amazon.com logo is displayed in white text on an orange background. The word "amazon" is in a lowercase, sans-serif font, followed by ".com" in a smaller font. A curved orange arrow is positioned below the "a" and "z", pointing from the "a" to the "z".

Video Games

Summary

Amazon is interested in building a customized recommendation system based on customer's ratings on **video games**.

- Customized recommendation vs. general recommendation
- How to trim down a long list of recommended video games

Outline

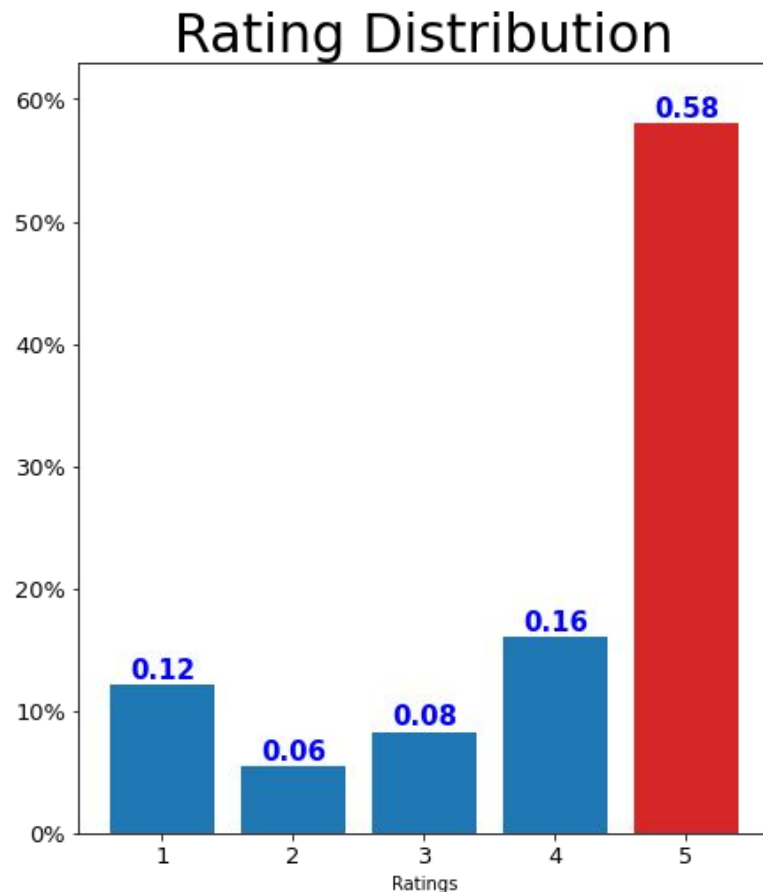
- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- 35% of Amazon's revenue from recommendations system⁽¹⁾
- Various recommendations: *popularity, purchase history, browsing history*
- No recommendation based on **customers ratings**
- Video games market size increases every year⁽²⁾
- Build a new recommendation system for **video games**
- Can the new system give customers a **different** shopping experience?
- Any chance for increased sale?

Data

- Reviews on **Video Games**⁽³⁾
- Two datasets
 - Years between 1996 and 2018
- Rating Data
 - 2.5 million ratings scaling from 1 to 5
 - 1.5 million customers
 - 72k **video games**
- Meta Data
 - 72k **video games**
 - Title, game id number, category, brand



Methods 1 - Build a System

1. Calculate ratings

2. Rank

3. Recommend

4. Test

- Existing rating vs. calculated rating
- Error range = ± 1.27**
- Lower error

	User 1	User 2	User 3	User 4
Item 1	5	?	3	?
Item 2	4	5	?	?
Item 3	?	3	2	?
Item 4	2	5	3	?
Item 5	?	1	?	5

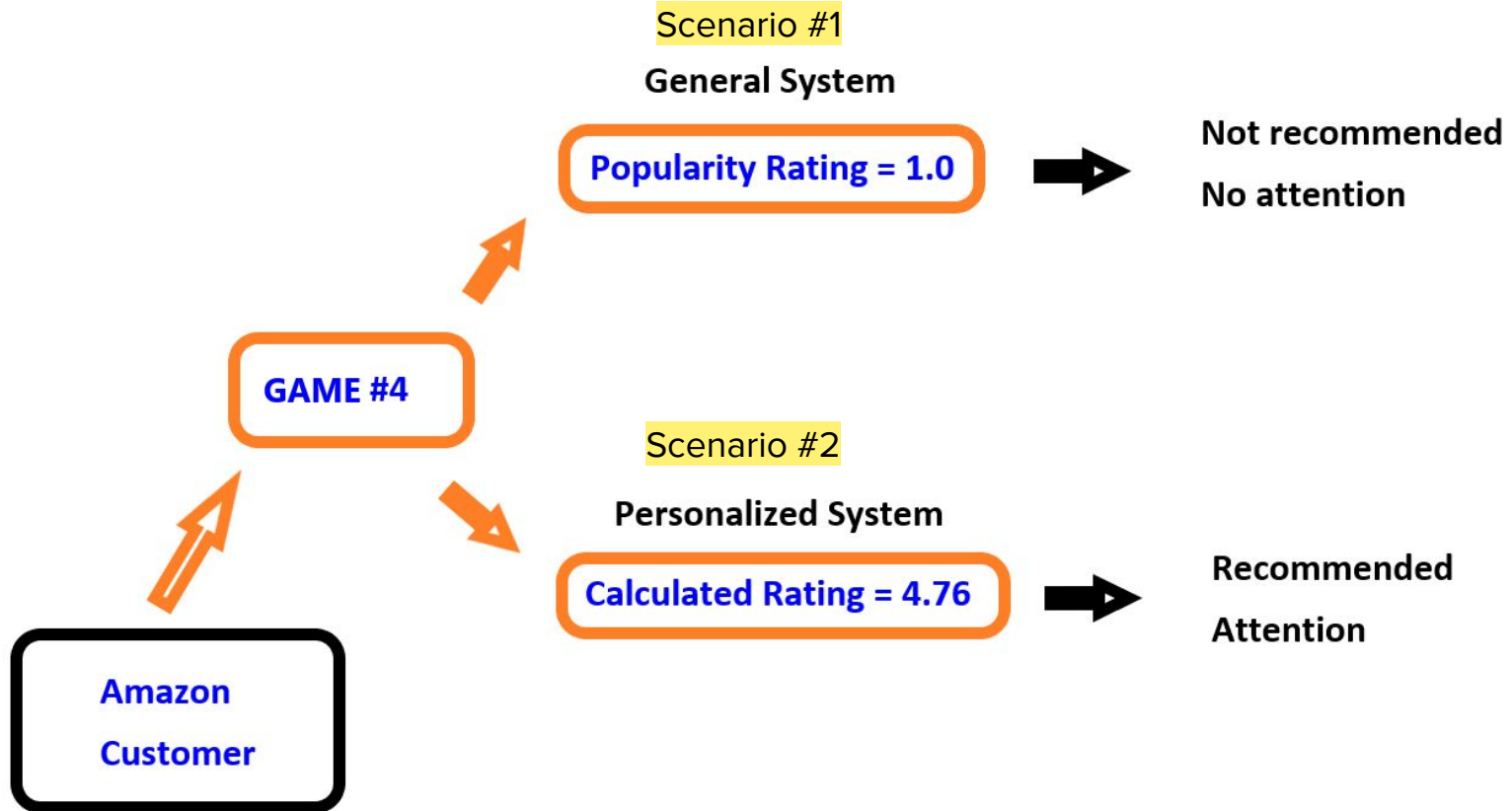
Methods 2 - Checking Difference on a Single Item



- Game #4 with ratings of **1.0 vs 4.76**
- Generally not recommended
- Recommended by personalized system
- **Different value**
 - Different shopping experience
 - Increased chance to buy Game #4

Video Game	Popularity Rating	Calculated Rating
Game 1	5	4.76
Game 2	4.75	4.91
Game 3	5	4.83
Game 4	1	4.76
Game 5	3	4.76

Methods 2 - Customer Experience Diagram



Methods 2 - Checking Difference on Multiple Items

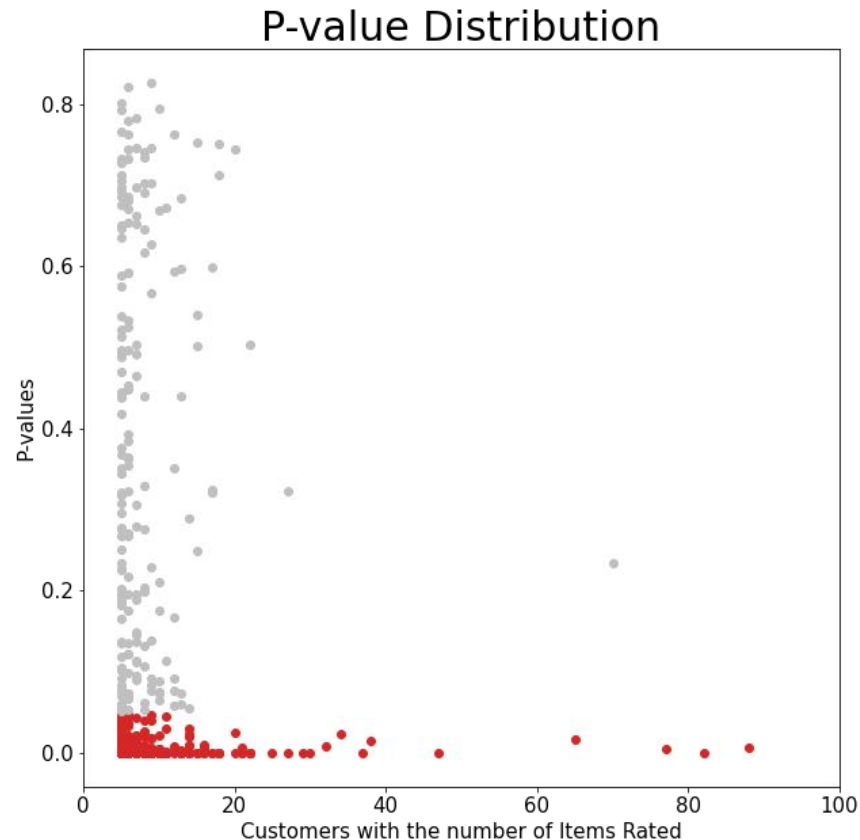


- Select a customer
- Assign calculated rating
- Comparison on two rating types
- Check p-value
 - Ranges from 0 to 1
 - Lower than 0.05 - significant difference
- **Assign p-value to the customer**
- Repeat on different customers

Video Game	Popularity Rating	Calculated Rating
Game 1	5	4.76
Game 2	4.75	4.91
Game 3	5	4.83
Game 4	1	4.76
Game 5	3	4.76
...
Game #	3	3.86

Results - P-values

- 500 random customers selected
- **311 customers**
 - **p-value < 0.05**
 - **Different shopping experience**

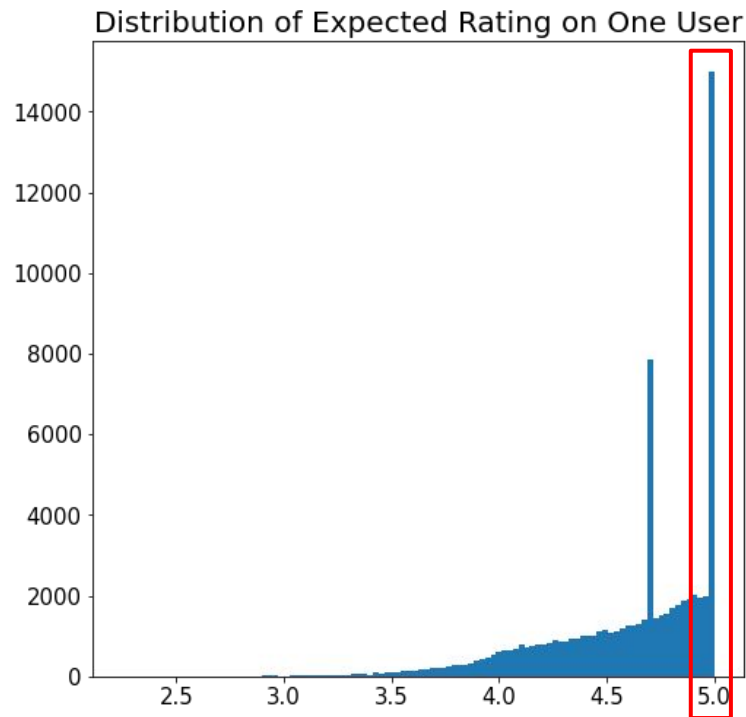


Methods 3 - Trimming Recommendations

- More than 15,000 w/ calculated rating ≈ 5

- **Similarity**

- Select one video game
 - Brand - Capcom, Blizzard, Sony
 - Console - PC, Xbox, Playstation, Nintendo
- Ranges from 0 to 1
- Multiply **similarity** to calculated ratings



Results - Trimming with Similarity

Resident Evil 7 Biohazard - Xbox one (selected game)

Similarity	
Perfect	15 games
High	0.5%
Good	6%
Mid	10%
None	84%

Recommended Games
DMC Devil May Cry: Definitive Edition - Xbox One
Resident Evil 6 - Xbox One
Strider - Xbox One Digital Code
Devil May Cry 4: Special Edition - Xbox One
Resident Evil Origins Collection - Xbox One Standard Edition

Conclusions

- More than 60% of customers are expected to get a different shopping experience from the personalized recommendation system built based on customers' ratings on video games. We can expect an increase in sales of games with low popularity ratings.
- A list of many games with calculated ratings ≈ 5 can be trimmed using the [similarity](#) on one game.

Next Steps

- Lower the error range
 - Control data size
 - More recent data
 - Set minimum number of ratings per customer
- Gather more information for improved trimming
 - Genre (action, shooting, adventure, etc)
 - Release Date
- Discuss on where and how to display recommendations
- Repeat same process with items from a different department

Thank You!

Email: youremail@email.com

GitHub: @username

LinkedIn: [linkedin.com/in/username/](https://www.linkedin.com/in/username/)

Sources

1. <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers>
2. <https://www.ibisworld.com/industry-statistics/market-size/video-games-united-states/>
3. <https://nijianmo.github.io/amazon/index.html>