Recommender System Video Games

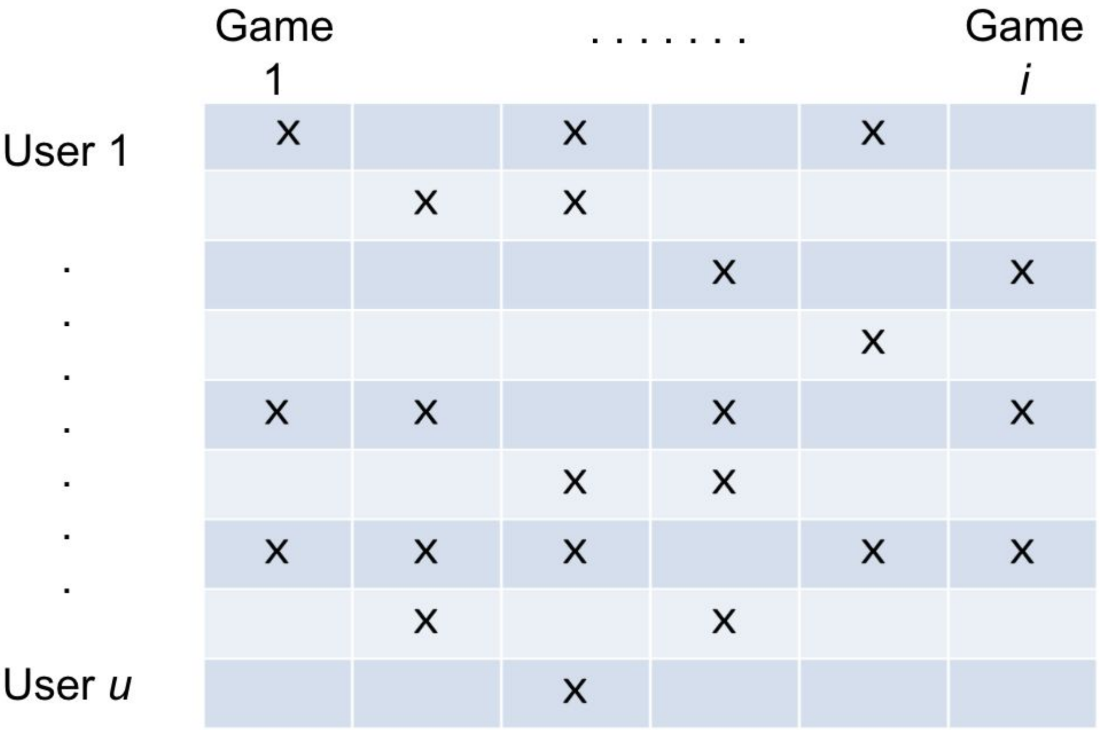
Sam Darmali

The Business Problem

- Companies have large catalogues of products.
- Impossible for customers to try every single product.
- Helps users discover products that they may not have otherwise found.
- Leads to more conversions and increased revenue.

Data Overview

- Amazon Customer Reviews Dataset
  - Ratings: 1,700,000+
  - Unique users: 1,000,000+
  - Unique games: 56,000+
  - Sparsity: **0.003%**

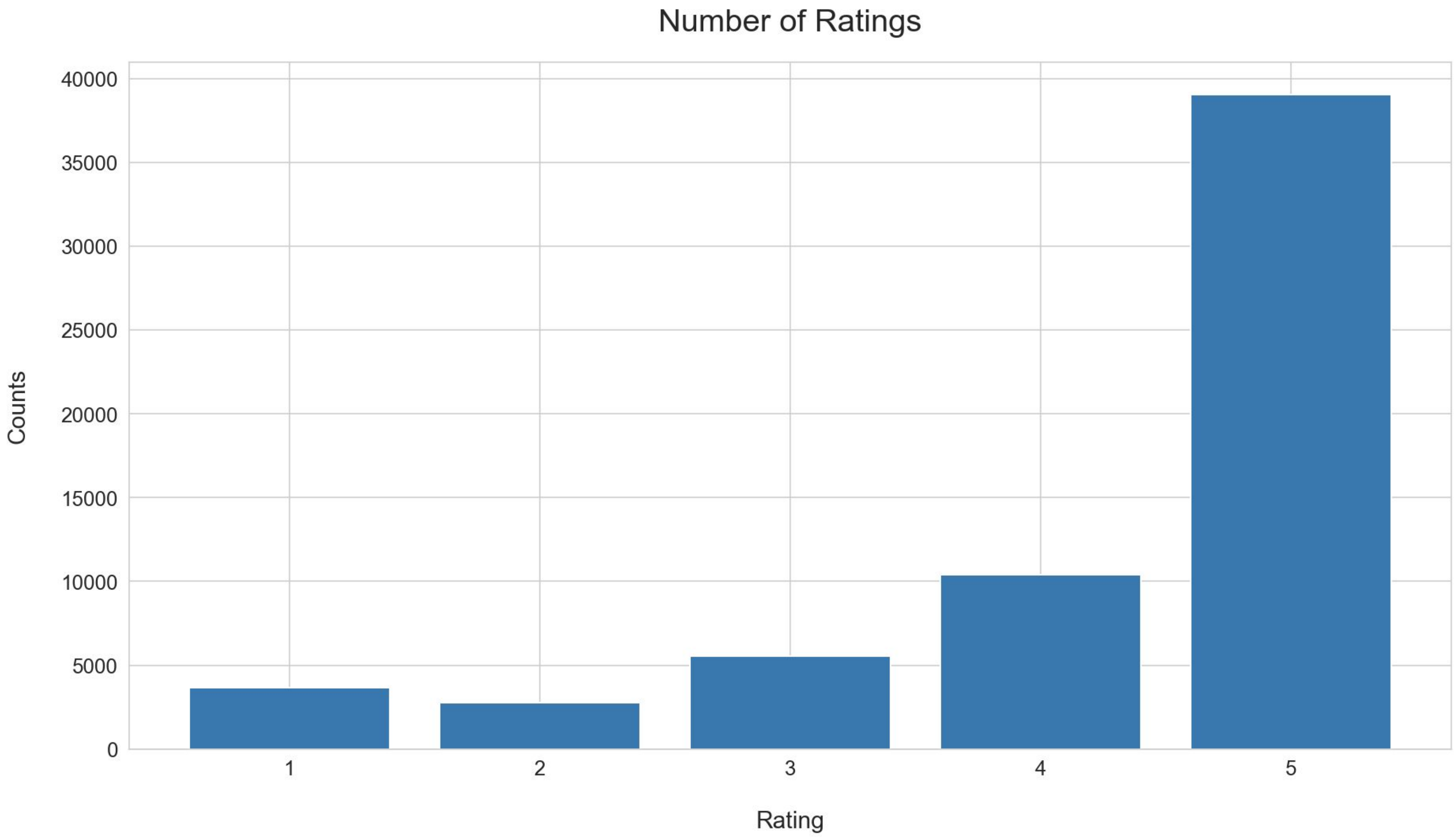


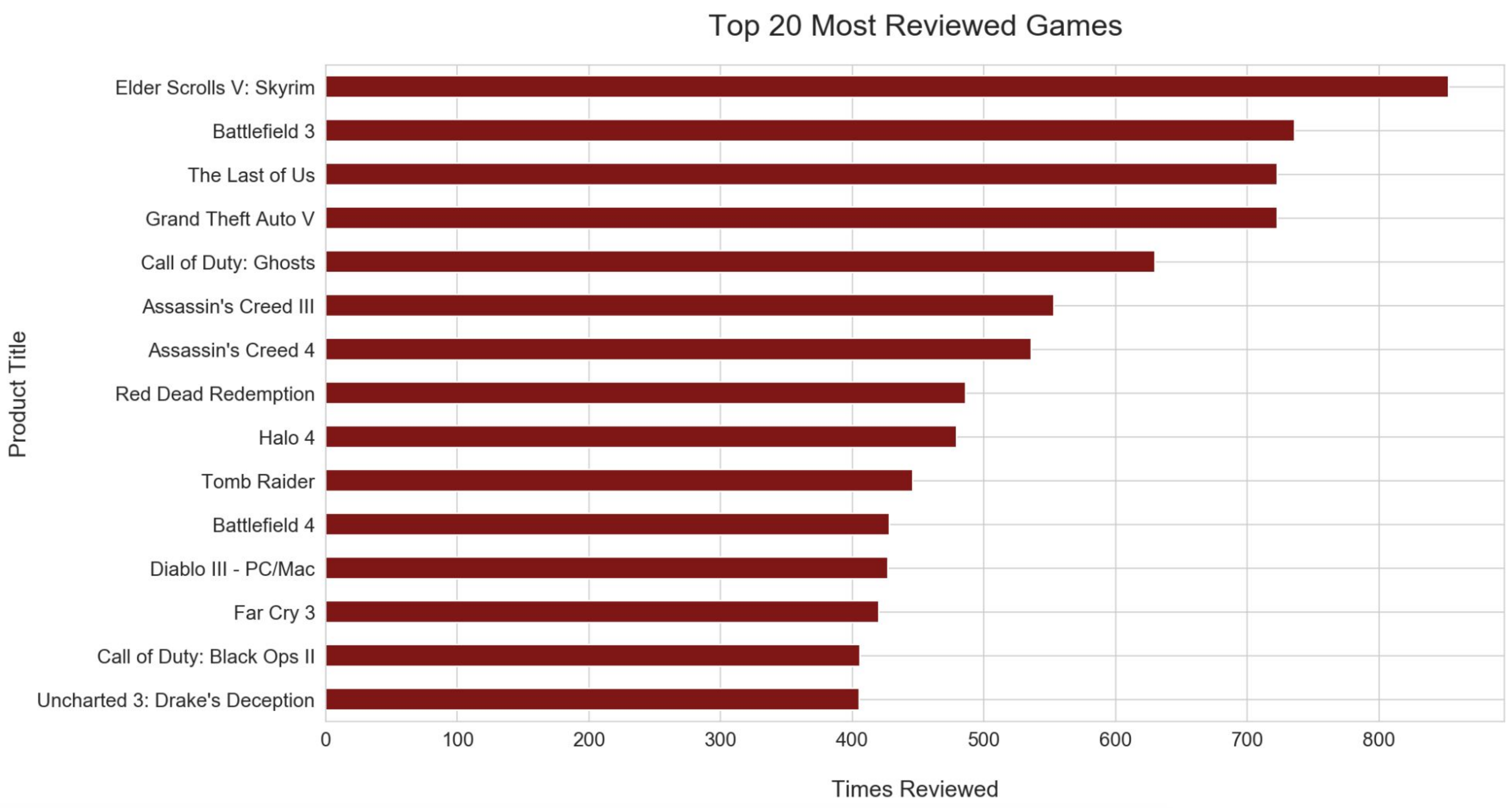
Subsetting the Data

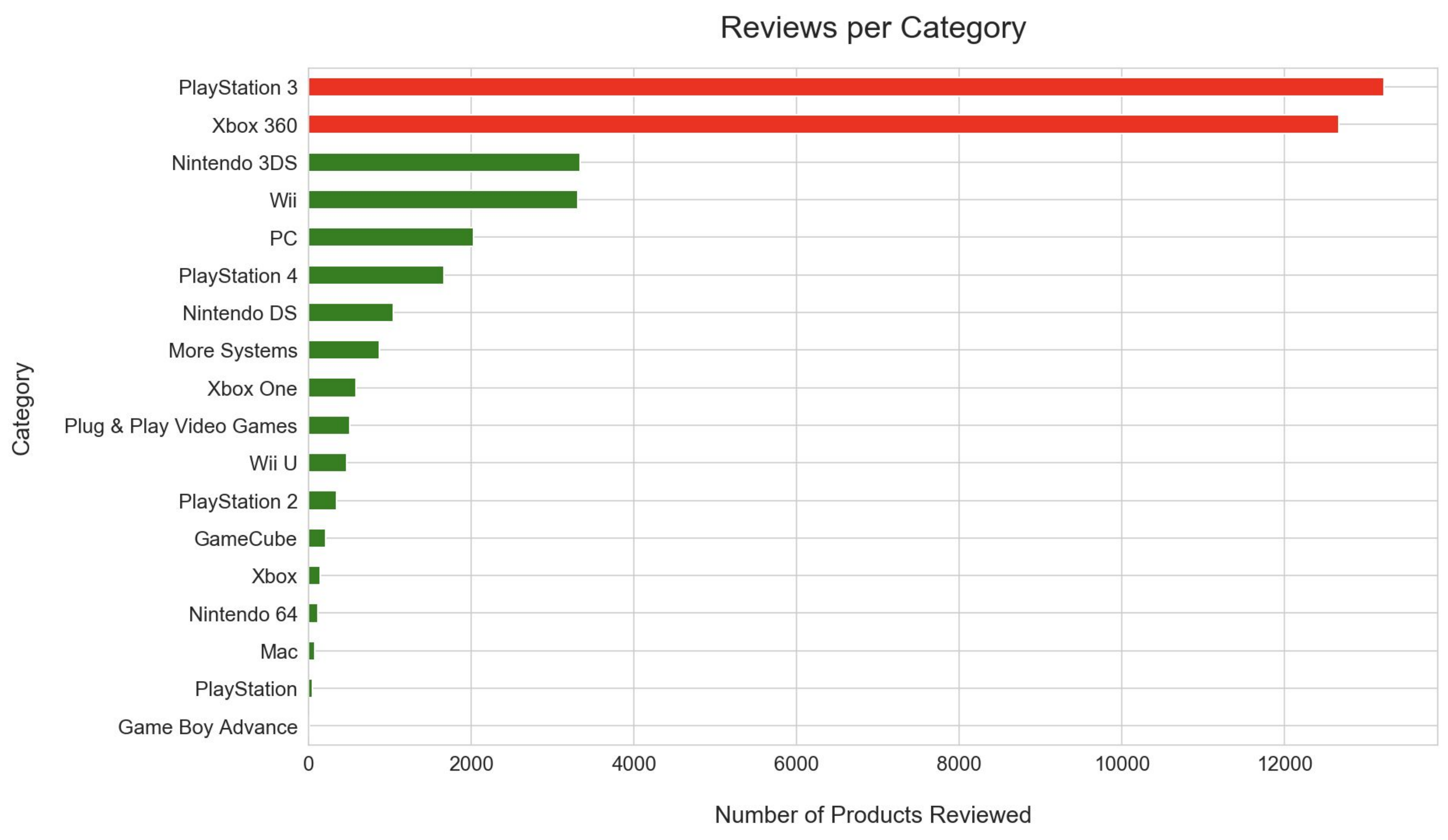
- The casual gamer
    - Has reviewed more than 5 games
    - Has not reviewed more than 300 games (Youtubers / professional reviewers)
- Games liked by the casual gamer
    - Has received more than 400 reviews
- Time frame
    - Ratings from 2010 - 2015
    - Train: 2010 - 2013
    - Test: 2014 - 2015

Data Subset

- Subset:
    - Ratings: 33,000+
    - Unique users: 13,000+
    - Unique games: 200+
    - Sparsity: **1.126%**


Number of Ratings
Counts
Rating
0
5000
10000
15000
20000
25000
30000
35000
40000
1
2
3
4
5

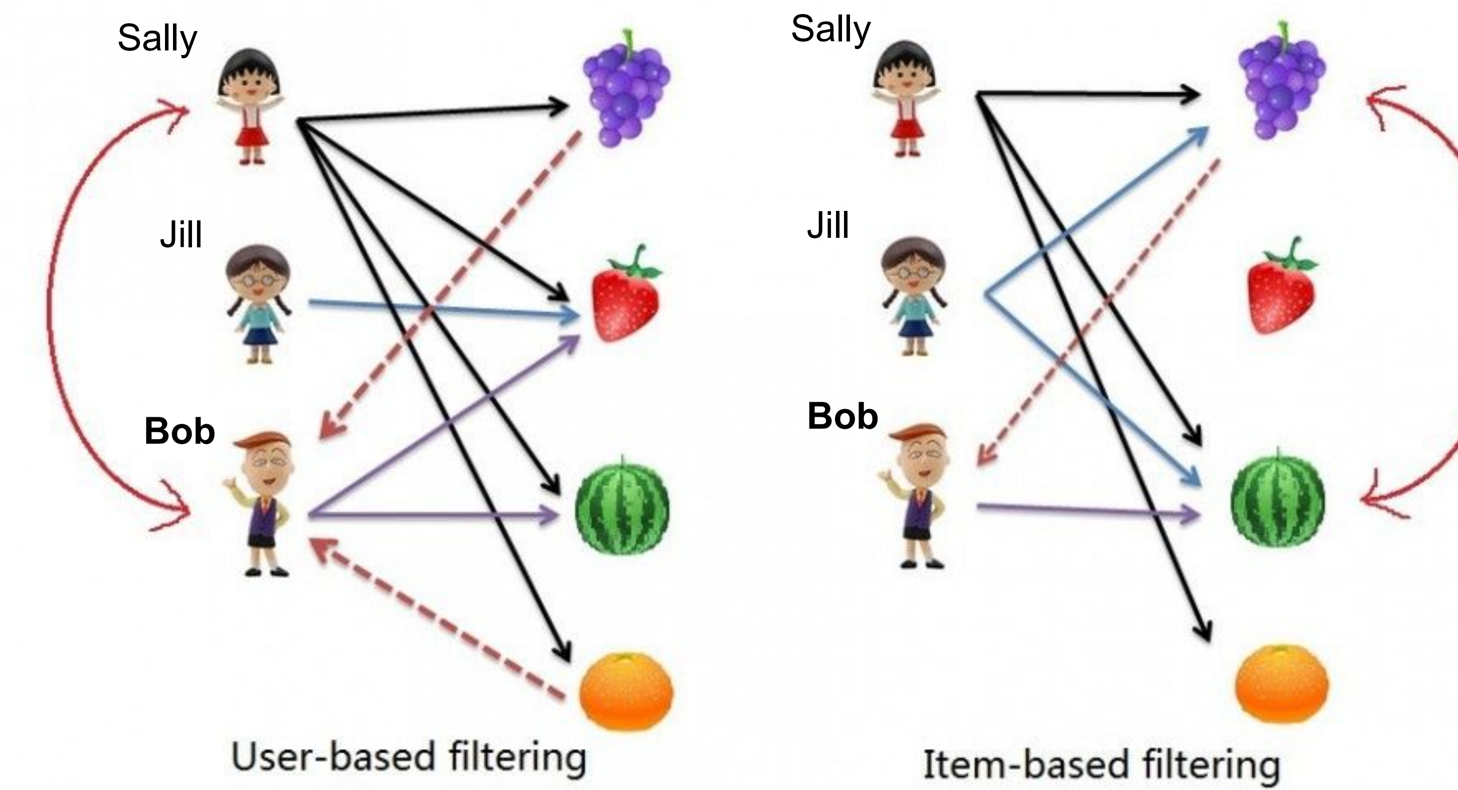

Top 20 Most Reviewed Games
Product Title
Elder Scrolls V: Skyrim
Battlefield 3
The Last of Us
Grand Theft Auto V
Call of Duty: Ghosts
Assassin's Creed III
Assassin's Creed 4
Red Dead Redemption
Halo 4
Tomb Raider
Battlefield 4
Diablo III - PC/Mac
Far Cry 3
Call of Duty: Black Ops II
Uncharted 3: Drake's Deception
0
100
200
300
400
500
600
700
800
Times Reviewed


Reviews per Category
PlayStation 3
Xbox 360
Nintendo 3DS
Wii
PC
PlayStation 4
Nintendo DS
More Systems
Xbox One
Plug & Play Video Games
Wii U
PlayStation 2
GameCube
Xbox
Nintendo 64
Mac
PlayStation
Game Boy Advance
Category
0
2000
4000
6000
8000
10000
12000
Number of Products Reviewed

Approach

- Modelling
  - User-based collaborative filtering
  - Item-based collaborative filtering
  - Latent factor models (Matrix Factorization)
- Evaluation
  - Choose 3 best models based on:
    - Root Mean Squared Error
    - Mean Precision
    - Mean Average Precision

Collaborative Filtering



Matrix Factorization

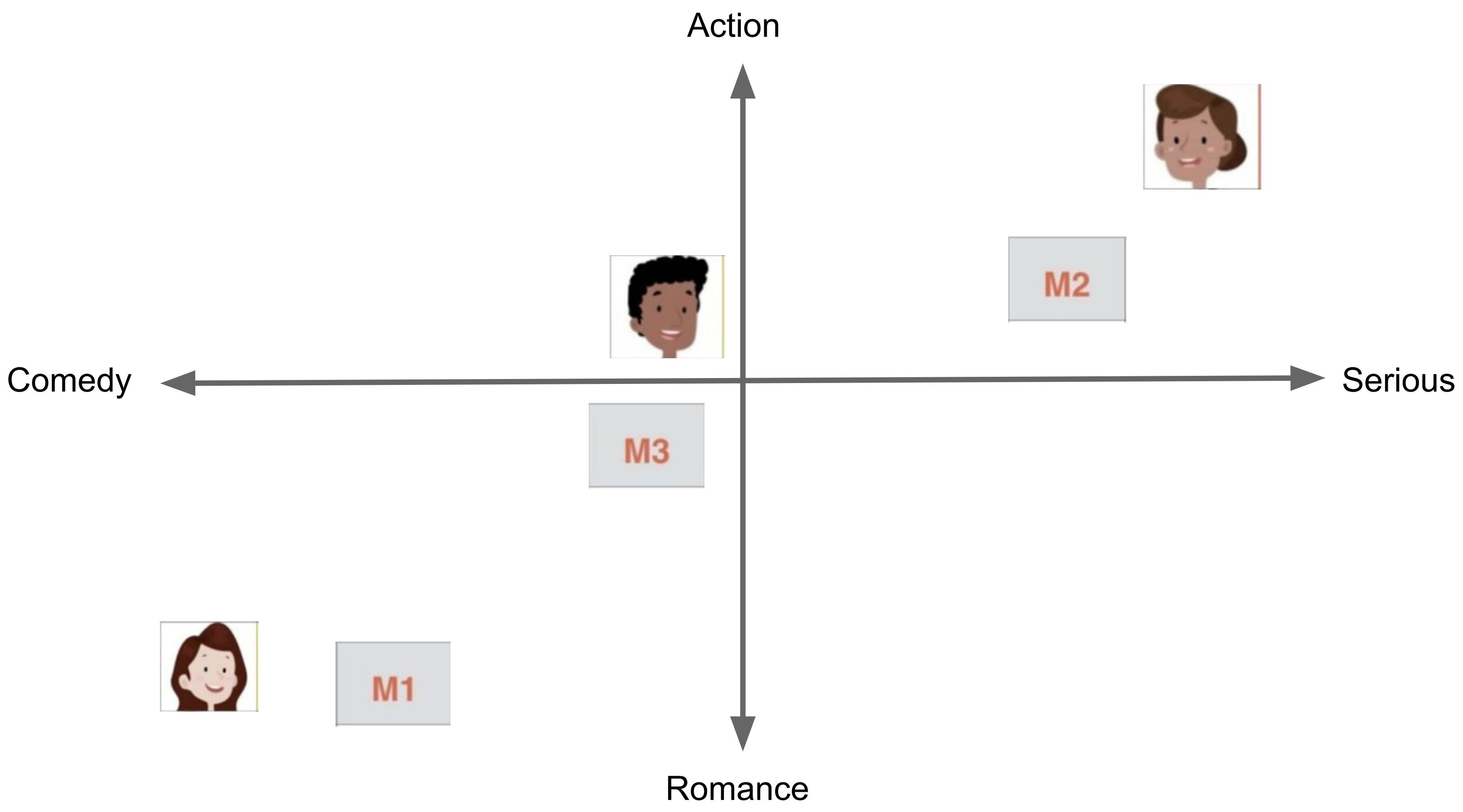
- Characterizes items and users by vectors of latent factors

	M1	M2	M3	M4	M5
Comedy	3	1	1	3	1
Action	1	2	4	1	3

- Latent factors inferred from item rating patterns

	Comedy	Action
A	✓	✗
B	✗	✓
C	✓	✗
D	✓	✓

	M1	M2	M3	M4	M5
	3	1	1	3	1
	1	2	4	1	3
	3	1	1	3	1
	4	3	5	4	4


Action
M2
Comedy
Serious
M3
M1
Romance

Success Criteria / Metrics

- Root Mean Squared Error (RMSE)
    - Measure of how close predicted ratings are to actual ratings.
    - Problem:
        - Predicts 5 stars for many games
        - Poor at choosing which of those to recommend

Success Criteria / Metrics

- Mean Precision

$$Precision = \frac{\text{\# of our recommendations that are relevant}}{\text{\# of items we recommended}}$$

  - 'Relevant' = Rated 4 or above

- Mean Average Precision (MAP)
  - Average Precision (AP) - Average of Precision scores up to a given cutoff.

$$\text{AP@N} = \frac{1}{m}\sum_{k=1}^{N}(P(k) \text{ if } k^{th} \text{ item was relevant}) = \frac{1}{m}\sum_{k=1}^{N} P(k) \cdot rel(k)$$

$$\text{MAP@N} = \frac{1}{|U|}\sum_{u=1}^{U}(\text{AP@N})_u = \frac{1}{|U|}\sum_{u=1}^{U}\frac{1}{m}\sum_{k=1}^{N} P_u(k) \cdot rel_u(k)$$

Mean Average Precision Intuition

Rank	Recommendation	Result	Precision
1	Call of Duty	True Positive	1
2	Grand Theft Auto	False Positive	1/2
3	Assassin's Creed	False Positive	1/3

$$AP = \frac{1 + 1/2 + 1/3}{3}$$

$$= 0.61$$

Rank	Recommendation	Result	Precision
1	**Assassin's Creed**	False Positive	0
2	Grand Theft Auto	False Positive	0
3	**Call of Duty**	True Positive	1/3

$$AP = \frac{0 + 0 + 1/3}{3}$$

$$= 0.11$$

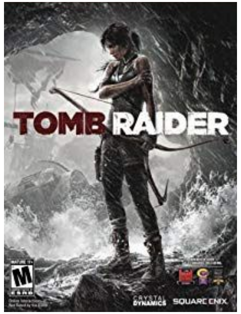
Quantitative Findings

	Root Mean Squared Error	Mean Precision	Mean Average Precision
Matrix Factorization Model	1.09	0.85%	8.22%
User Based Model	1.33	1.50%	16.03%
Item Based Model	1.21	1.47%	16.31%

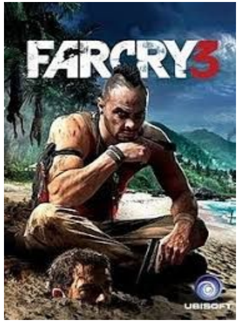
Genres

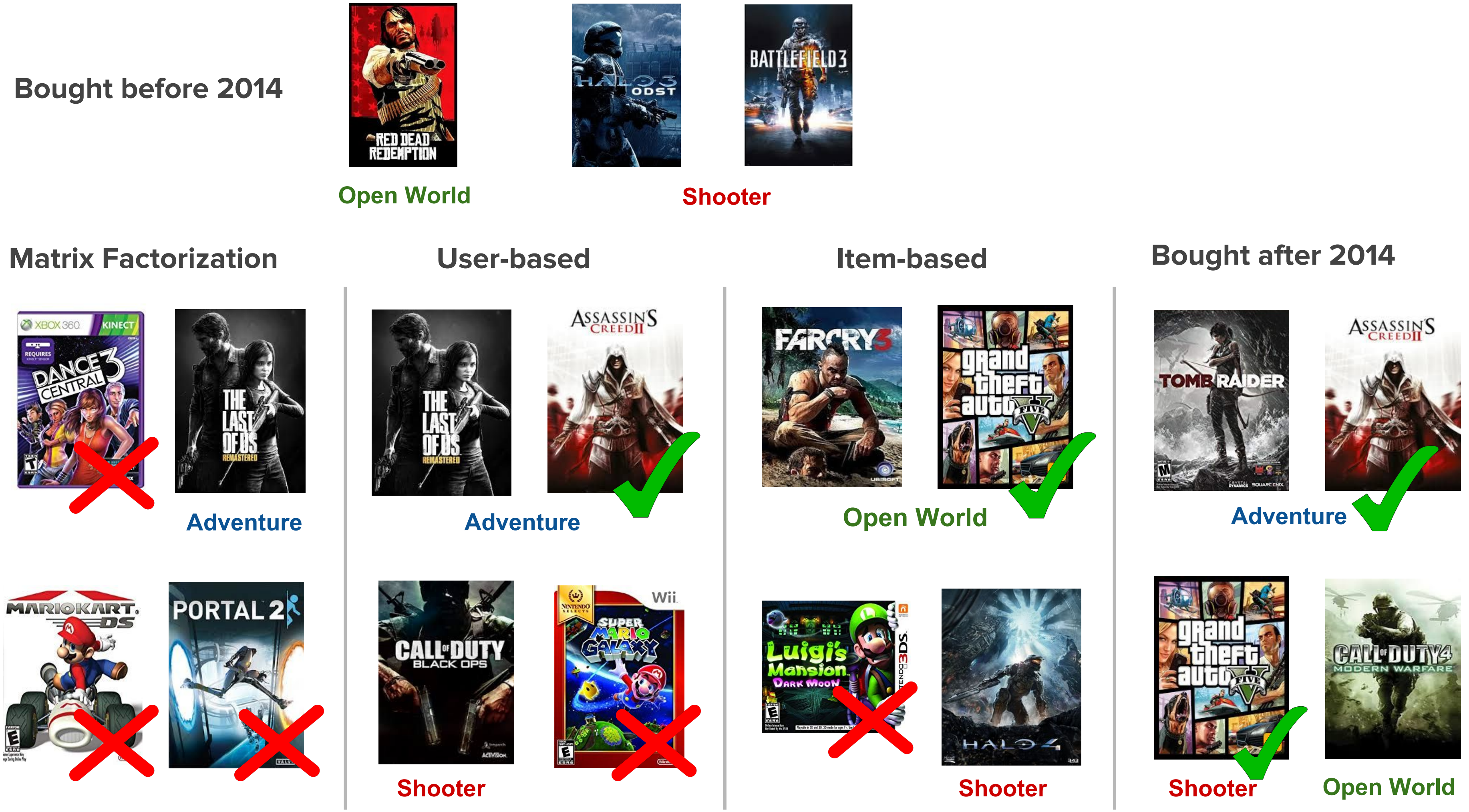
Shooters

Adventure



Open World




Bought before 2014
RED DEAD REDEMPTION
HALO 3 ODST
BATTLEFIELD 3
Open World
Shooter
Matrix Factorization
User-based
Item-based
Bought after 2014
DANCE CENTRAL 3
THE LAST OF US REMASTERED
Adventure
MARIOKART DS
PORTAL 2
THE LAST OF US REMASTERED
ASSASSIN'S CREED II
Adventure
CALL OF DUTY BLACK OPS
SUPER MARIO GALAXY
Shooter
FARCRY 3
grand theft auto FIVE
Open World
Luigi's Mansion Dark Moon
HALO 4
Shooter
TOMB RAIDER
ASSASSIN'S CREED II
Adventure
grand theft auto FIVE
CALL OF DUTY 4 MODERN WARFARE
Shooter
Open World

Limitations

- Ratings are explicit feedback
- Implicit feedback would be helpful to gauge interest
    - Browsing history
    - Search patterns
    - Mouse movements

Conclusions

- Lessons
  - The business case defines the success metric
  - For marketplace recommenders, precision / ranking based metrics are preferred
- Next
  - Fix Mario problem
  - Deployment
  - Content based filtering
  - Deep learning

Thank You