



Amazon Ratings & Review Dataset

# VIDEO GAMES



Recommendation Systems  
Group E - GMBD 2021

Ramon Denia | Nalisha Men | Jan P. Thoma | Elias Theodoropoulos | Disha Saxena



# BACKGROUND

a solution to the paradox of choice among Video Games lovers

Building a product recommendation system scenario for video games and surrounding products listed on Amazon. Implementing collaborative filtering and content-based models to find suitable items for recommendation.

Overall, the video game consumer prefers a game genre (Shooter and Actions are two leading genres), and often purchases new games within the same game console, which is a highly specialized device. The business evaluation follows these two aspects.

Data Source

Jianmo Ni, Jiacheng Li, Julian McAuley

Empirical Methods in Natural Language Processing (EMNLP), 2019

<https://nijianmo.github.io/amazon/index.html>



# COLLABORATIVE FILTERING

KNN and SVD from Surprise

Based on user overall ratings (1-5) per items, an item-to-item model is built.



# MODELING

Distribution Of 1521260 Video Games ratings



```
{'sim_options': {'name': 'msd', 'min_support': 3, 'user_based': False}}
```

```
cross_validate(svdtuned, data, measures=['RMSE', 'MAE'], cv=3, verbose=True)
```

Evaluating RMSE, MAE of algorithm SVD on 3 split(s).

	Fold 1	Fold 2	Fold 3	Mean	Std
RMSE (testset)	1.3305	1.3306	1.3211	1.3274	0.0044
MAE (testset)	1.0423	1.0447	1.0360	1.0410	0.0037
Fit time	12.55	11.93	14.55	13.01	1.12
Test time	0.13	0.25	0.16	0.18	0.05

```
{'test_rmse': array([1.33047326, 1.3306096 , 1.32110339]),  
 'test_mae': array([1.04234332, 1.04473054, 1.03598346]),  
 'fit_time': (12.548882961273193, 11.93193793296814, 14.55277705192566),  
 'test_time': (0.12697196006774902, 0.25313377380371094, 0.16203284263610  
 84)}
```

## Sampling

Stratified sampling of the original dataset to ensure ratings (from 1 to 5) are sampled proportionally within the new dataset of 60000 (~4% of original dataset).

## Benchmark

KNNBaseline and SVD seem to have better benchmark performance.

Algorithm	test_rmse
KNNBaseline	1.328286
BaselineOnly	1.328521
SVD	1.328849
SVDpp	1.329433
SlopeOne	1.361816
CoClustering	1.362069
NMF	1.370546
NormalPredictor	1.726713

## Optimize parameters: KNNBaseline & SVD

Parameters are further tuned for two models and cross validate to evaluate the performance.

10 nearest neighbors of

# NINTENDO JUST DANCE



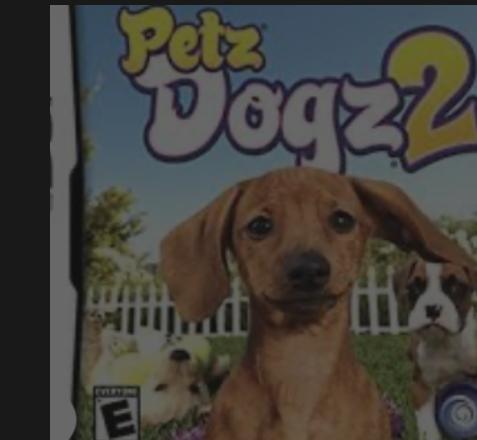
10 nearest neighbors of Just Dance 2014 – Nintendo Wii are:

- Spore Hero – Nintendo Wii
- Xbox One Media Remote
- PlayStation 4 Black Friday Bundle – Grand Theft Auto V and The Last of Us Remastered
- Skylanders Trap Team, Micro Comic Fun Packs, Jawbreaker with Exclusive Undead Trap
- Xbox One Chat Headset
- Dreamgear 3-Piece Stylus Pack for Nintendo 3DS
- Ratchet & Clank Collection
- Xbox 360 Universal Media Remote
- Wii Fit U w/Fit Meter – Wii U
- Dogz 2



## ✓ Nintendo Wii Games

Spore-Hero  
Dogs 2  
Skylander



## ✓ Wii- Gear related

Wii Fit U w/Fit Meter - Wii U  
Dreamgear 3-Piece Stylus Pack



## Games in Other Consoles

Xbox  
PlayStation 4

10 nearest neighbors of

# XBOX 360 LEGO MOVIE



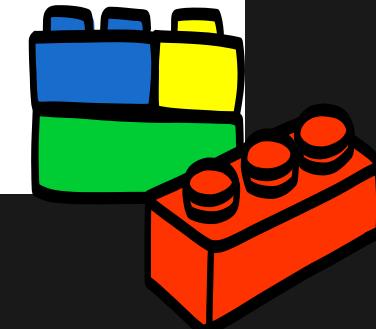
## ✓ Xbox-related Games

- Madden NFL 25
- FIFA 14
- Gears of War
- Call of Duty 4
- Batman Returns
- Disney: Star Wars Rebels Zeb Orrelios Figure



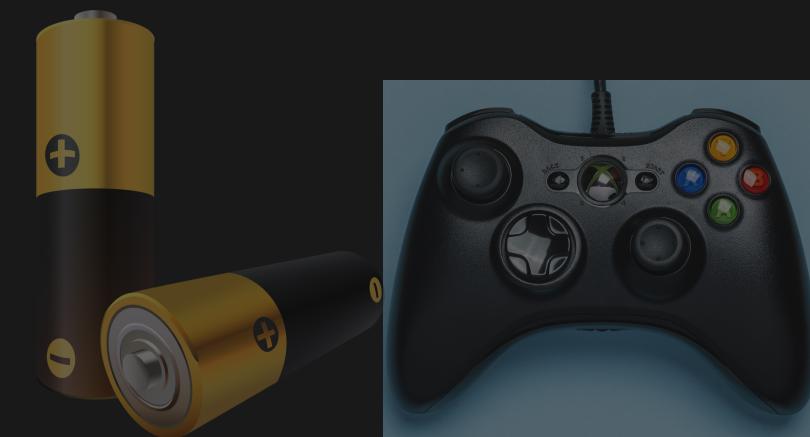
The 10 nearest neighbors of The LEGO Movie Videogame – Xbox 360 Standard Edition are:

- Xbox 360 Rechargeable Controller Battery Pack White
- Watch Dogs – PlayStation 4
- Batman Returns
- Madden NFL 25 – Xbox 360
- FIFA 14 – Xbox 360
- Disney Infinity 3.0 Edition: Star Wars Rebels Zeb Orrelios Figure
- F1 2016 – PlayStation 4
- Call of Duty 4: Modern Warfare [Download]
- Gears of War: Ultimate Edition – Xbox One
- Stella Glow – Nintendo 3DS



## ✓ Device Related

- Rechargeable Controller Battery Pack

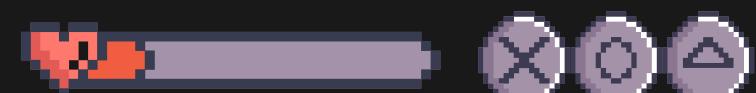


## Games in Other Consoles

- Nintendo
- PlayStation 4



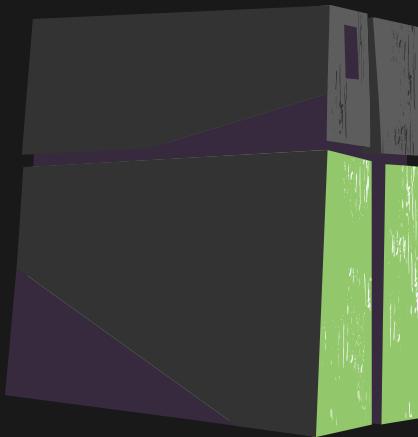
10 nearest neighbors of



# THE LAST OF US

## + Playstation-related Games

- Madden NFL 25
- The Elder Scrolls V
- Disney Infinity: Marvel Super Heroes



## + Device Related

- PS4 Charging Station
- Protective Starter Kit
- Stick Cover

## Games in Other Consoles

Nintendo  
Xbox

The 10 nearest neighbors of The Last of Us Remastered – PlayStation 4 are:

Madden NFL 25 Anniversary Edition with NFL Sunday Ticket – Playstation 3  
The Elder Scrolls V: Skyrim Legendary Edition – Playstation 3  
ACC PS4 DUALSHOCK 4 CHARGING STATION BY SONY #MAIN  
Nintendo Nunchuk Controller – White  
HORI Elite Pack Protective Starter Kit for PlayStation Vita 2000  
Fallout 4 – Pip-Boy Edition – Xbox One  
Xbox One 500GB Console – Gears of War: Ultimate Edition Bundle  
Cyber & Analog Stick Cover High Type (For Ps Vita) White  
Disney INFINITY: Marvel Super Heroes (2.0 Edition) Power Disc Pack  
Nintendo 3DS Compatible with 3DS / 3DS XL / 2DS AC Adapter



# OBSERVATION (I)

## COLLABORATIVE FILTERING



### **Game-console sensitive**

The model has captured the customers' behavior that software (games) are purchased according to the console of the hardware unit.



### **Hardware / Game-gears Specific**

Fit-gears (Wii-related IoT devices) are captured by the model for the item Just Dance (Nintendo-Wii). While for Xbox games, it is reasonable that most device interaction happens when buying rechargeable battery or replacing a broken controller.



### **Game Genres are not clear**

Without a clear game genre marker (Action, Casual, Shooter etc.), it is inevitable to have Call of Duty recommended for Lego Movie Videogame gamers, and NFL for Last of Us within the same game console.



# CONTENT-BASED RECOMMENDER SYSTEM

TF-IDF Representation and Cosine Similarity

Leverage the textual information of item description to link items



# MODELING

	000	10	100	12	13	15	16	20	2000	25	...	working	works	world	worlds	xbox	xl	xp	year
0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.199878	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.183035	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
45704	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
45705	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
45706	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0
45707	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.079824	0.0	0.000000	0.0	0.0	0.0
45708	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	...	0.0	0.0	0.000000	0.0	0.579299	0.0	0.0	0.0

```
{'works': 0.7099328657881848,  
'used': 0.5150031172668346,  
'great': 0.48038225953839947,  
'10': 0.0,  
'platforms': 0.0,  
'playing': 0.0,  
'players': 0.0,  
'player': 0.0,  
'played': 0.0,  
'playable': 0.0,  
'play': 0.0,  
'platform': 0.0,  
'playstation': 0.0,  
'plastic': 0.0,  
'planet': 0.0,  
'places': 0.0,  
'place': 0.0,  
'pirates': 0.0,  
'pieces': 0.0,  
'plays': 0.0,  
'plug': 0.0,  
'plenty': 0.0,  
'pick': 0.0,  
'practice': 0.0,  
'powers': 0.0,  
'powerful': 0.0,  
'powered': 0.0,  
'power': 0.0,  
'possible': 0.0,  
'possibilities': 0.0,
```

## ►► Text Cleaning

Item descriptions contain strings in html format and unicode, such as "/xa0". Data cleaning to ensure texts are cleaned, and also are simplified through stemming method when possible.

## ►► Profiling Game items

Sampling ~3% of original data. Identify the most relevant 'words' in the textual descriptions for each item, and relevance is assigned with a score. Feature extraction is done via TF-IDF.

## ►► Linking items based on Cosine Similarity

Retrieve similar items based on cosine similarity method.

# NINTENDO JUST DANCE



✓ Dance-related Games

✓ Dance Gear related

Recommending 5 video games similar to Just Dance 2016...

- 
- Wii Dancing Mat (score:0.5075848268739979)
- Dance Dance Revolution Hottest Party 2 with Dance Mat (score:0.486261476789795)
- Dance Dance Revolution 5th Mix [Japan Import] (score:0.4637932708575073)
- Alvin and the Chipmunks: Chipwrecked (score:0.4566698308729095)
- Playstation 2 DDR Dance Pad (score:0.39550216617994)

# PC LEGO STAR WARS



✓ Star-war related Games

Recommending 10 video games similar to LEGO Star Wars II: The Original Trilogy (PC Games)

- 
- Star Wars Battlefront II (score:0.6805552776258814)
- Star Fox (Japanese Import) (score:0.5284104051248257)
- Star Wars: Empire at War (score:0.5116197629160257)
- Star Wars: The Clone Wars (score:0.48033337749692906)
- Star Wars Racer Revenge (score:0.3827528985837613)
- Wii Twin Star Sabers (score:0.3687356397438909)
- Lego Star Wars: The Complete Saga (score:0.29424260072101677)
- Nicktoons MLB 3D (score:0.24157006559672844)
- Star Trek Voyager: Elite Force (score:0.20554503708067168)
- Kirby: Nightmare in Dream Land (score:0.19654688773129)

# OBSERVATION (II)

## CONTENT-BASED



**Game-console sensitive**



**Hardware / Gear with related console**



**Repetitive content recommended**

Star war related games may lead to domination of all star-war games or items. No serendipities nor console-relevant games could be recommended.



**Game are better linked based on content**

Though not achieving stronger categorization based on typical games genre, such as Action, Casual, Shooter etc, items are better associated now from content perspective.

# RECOMMENDATIONS

## Optimize with clearer game genres

From the current dataset, we are unable to arrive at a clear game genre, such as Shooters, Action, Role-play, Strategy, Adventure and so on, based on Amazon item description or item category, though content-based model does provide a closer content association. Therefore, an enriched dataset with game genres would be useful to blend with the collaborative filtering model on user purchasing interaction, to achieve a console-specific and genre related game recommendation.



## Video Gamers as Amazon's consumers



It is critical to bear in mind the video gamers' unique customer behavior if they purchase on Amazon. Most often they are game-console specific and often loyal to a specific game or game genre. The CF model captures well the game console and can perform game-specific gear recommendation. However, for Amazon, to sell new games to these video game lovers might be extremely challenging without a "trial". A strong content-based model may lead to repetition and therefore loose the original incentive to recommend a "similar taste".