MyGamePass

Recommender System for Video Games

Ben Polzin

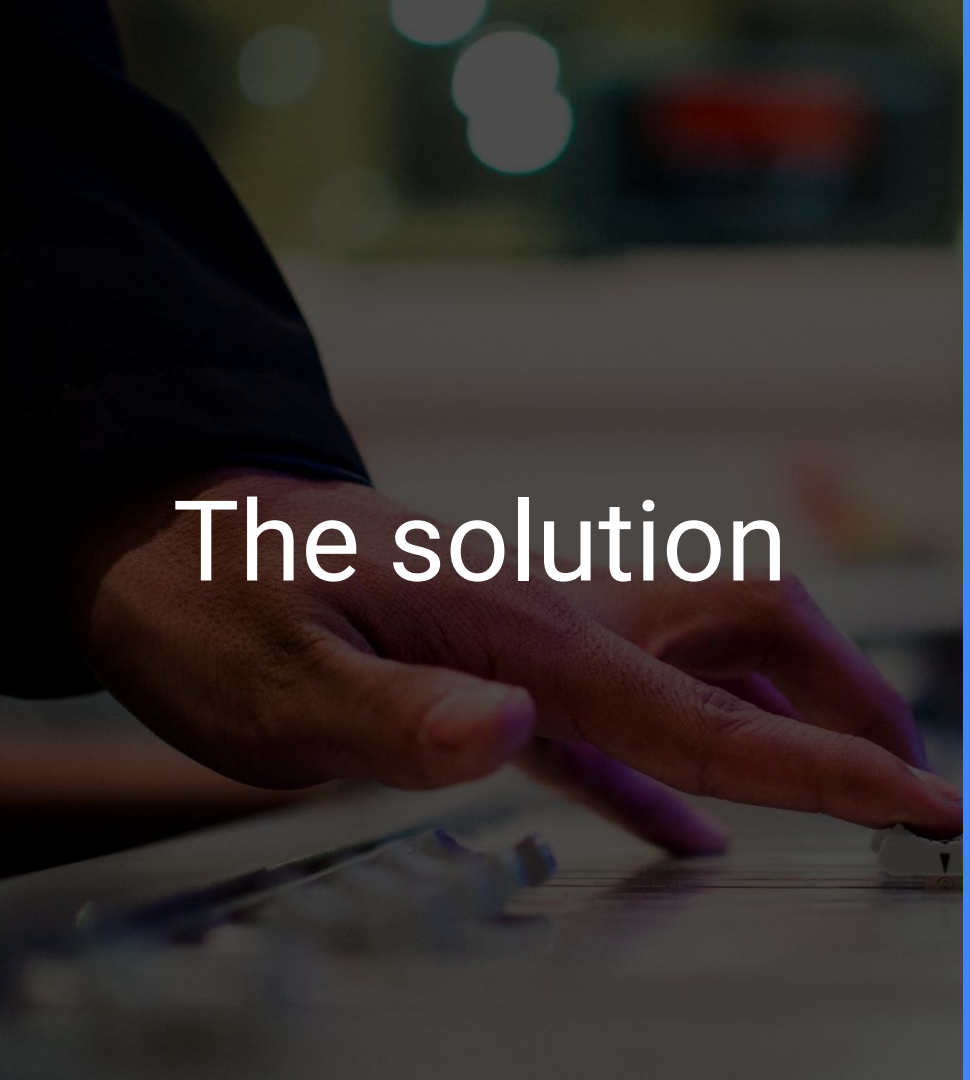
Data Scientist | B.S. Petroleum Engineering CSM ‘16

Passionate, Data Driven, and Future Focused

The Problem*: Too Many Games!

***so much variety, it is a good problem to have**

- ~1.2 Million video games across all platforms
  - PC, Console, Mobile
  - Action, Adventure, Puzzle, Platformer, etc.
- ~2.7 Billion individual games worldwide



The solution

Utilize modern machine learning techniques and natural language processing to develop a recommender system focused solely on the gamers and their games

How It Works
Content Based Filtering

Preprocessing

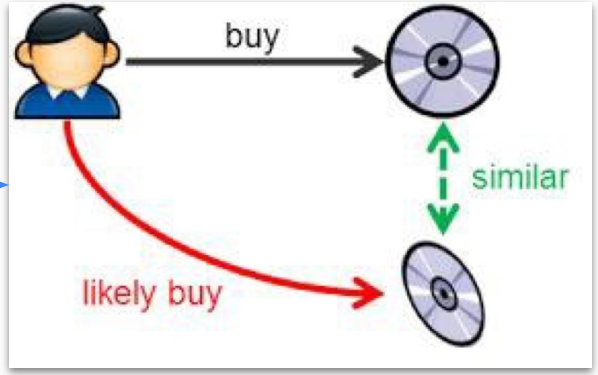
Clean and prepare game data for similarity matrix and process description with NLTK

ITEM-ITEM SIMILARITY MATRIX

	I_1	I_2		I_j		I_{m-1}	I_m
I_1	1	Sim_{12}	...	Sim_{1j}	...		
I_2		1	...		...		
	. . .	. . .	. . .	. . .	. . .	. . .	. . .
I_i		Simi2	...	Sim_{ij}	...		
	. . .	. . .	. . .	. . .	. . .	. . .	. . .
I_{m-1}			...		...	1	
I_m			...		...		1

Recommend Similar Content

Cosine similarity matrix is used to easily determine the most-similar items based on game content



Content Recommender Example


```
# Mass Effect Recommendation
similar_games2 = content_recommender('Mass Effect', games_df, similarities,
                                     vote_threshold=1000, rating_threshold=0.80)
similar_games2.head(5).sort_values(by='percent_positive_ratings',ascending=False)
```

	appid	game	similarity	vote_count	percent_positive_ratings
488	24980	Mass Effect 2	0.215164	11217	0.951680
143	6000	STAR WARS™ Republic Commando™	0.237015	6771	0.944026
382	17460	Mass Effect	1.000000	10773	0.938179
415	20540	Company of Heroes: Tales of Valor	0.165758	1438	0.929764
945	91200	Anomaly: Warzone Earth	0.180340	3781	0.872785

How It Works
Collaborative Based Filtering

Preprocessing

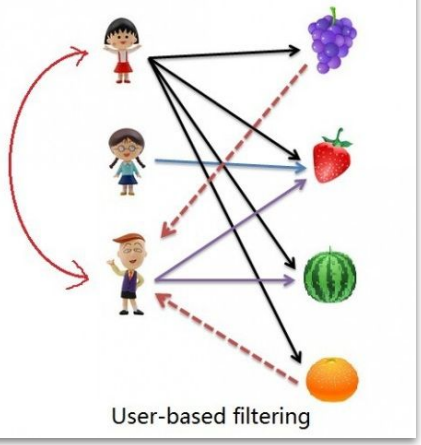
Clean and prepare user data for FunkSVD and Surprise package evaluations

USER-ITEM Matrix

	I_1	I_2		I_j		I_{m-1}	I_m
U_1			...		...		
U_2			...		...		
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
U_i			...	A_{ij}	...		
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
U_{n-1}			...		...		
U_n			...		...		

Recommend Similar Content

Similar user profiles are used to predict ratings for every game in the dataset to provide recommendations



Collaborative Recommender Example

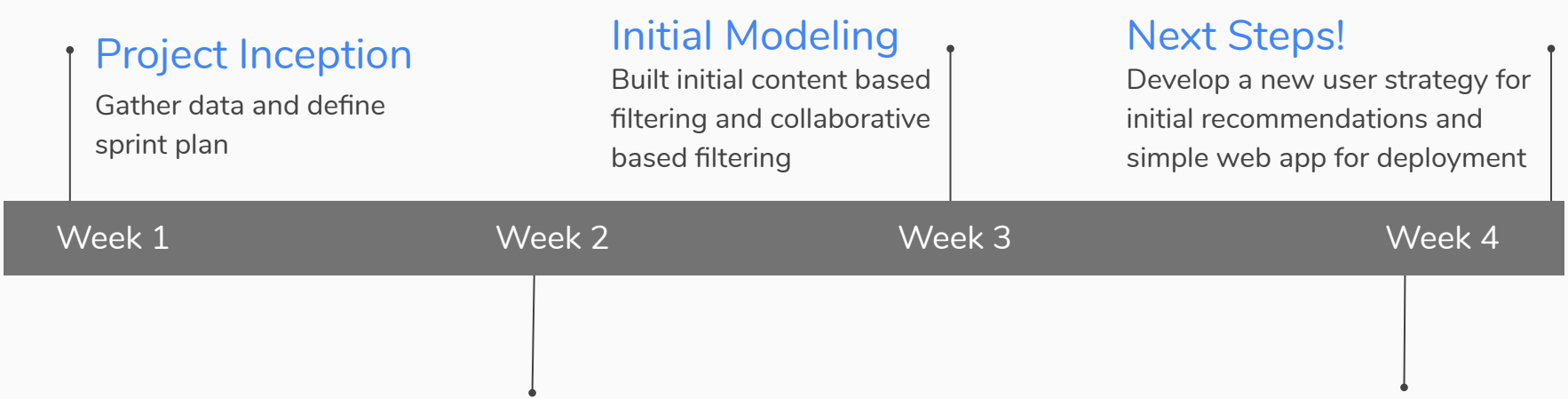

```python
# Check user id `128470551` top predictions
newdf = df_prediction[df_prediction['user_id'] == 128470551].sort_values(
    by=['predicted_rating'], ascending=False).head()

# Merge the dataframes
merge_df = newdf.merge(games_df, how='left',
                       left_on=['appid'], right_on=['appid'])
```

```
merge_df
```

predicted_rating	details	name	average_playtime	total_ratings	percent_positive_ratings
3.379807	{'was_impossible': False}	Torchlight II	300	31296	0.943859
2.432523	{'was_impossible': False}	Risk of Rain	300	24289	0.934579
2.347971	{'was_impossible': False}	Hammerwatch	288	5658	0.899965
1.994546	{'was_impossible': False}	Magic Duels	300	23511	0.706010

Milestones


Project Inception
Gather data and define sprint plan
Initial Modeling
Built initial content based filtering and collaborative based filtering
Next Steps!
Develop a new user strategy for initial recommendations and simple web app for deployment
Week 1
Week 2
Week 3
Week 4
Data Wrangling
Data cleaning, exploration and preprocessing for initial modeling
Model Optimization
Utilized FunkSVD matrix factorization as well as RMSE and FCP evaluations

Thank You!

- Huge thank you to the Education Team at BrainStation
- The awesome cohort for pushing me to constantly improve
- My Family

Ben Polzin

Data Scientist | B.S. Petroleum Engineering CSM ‘16

Passionate, Data Driven, and Future Focused

polzinben@gmail.com

linkedin.com/in/bpolzin

github.com/polzinben