

Movie Recommender

Xiaoxuan, Yiming, Xiaoyu, Violette

CONTENTS

1

Introduction

2

Simple
Recommender

3

Content-based
Recommender

4

Evaluations

5

Deployment

6

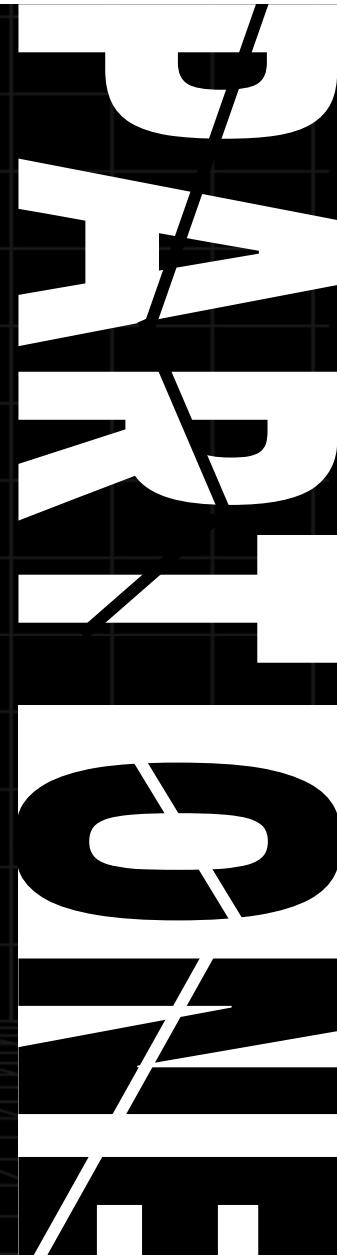
Future Plan



1

Introduction

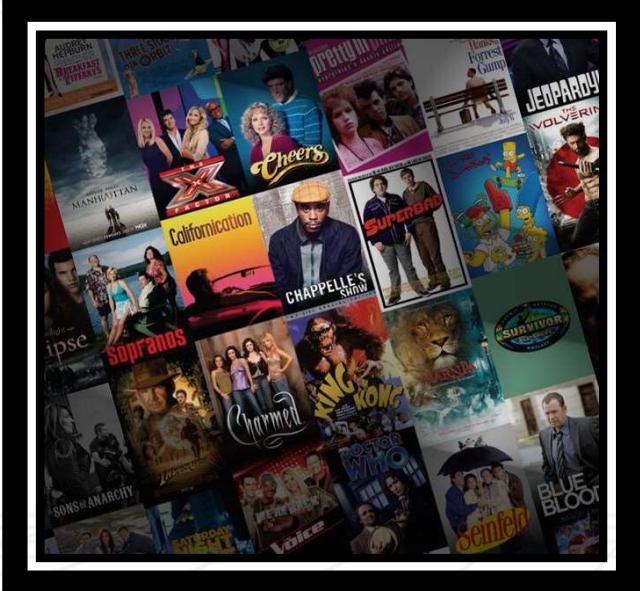
- Why?



The Movies Dataset

Kaggle

PART ONE



45,000
movies

26 million
ratings

From 1874
to 2017

Metadata:
cast, crew, plot keywords,
release dates, languages,
countries, TMDB votes



Simple Recommender

Generalized
Recommendations

2



The Simple Recommender

Ranking & Filter

PART TWO



Ranking

- Based on movie's **popularity**
 - solely based on vote count
- Based on movie's **rating**
 - weighted average on TMDB ratings from the dataset



Filters

genres,
languages,
year,
runtime,
director

The Simple Recommender

Ranking & Filter

PART TWO



```
# English action movies in 2012
```

```
build_full_chart("Action", "en", "2012", 0.85).head(10)
```

			title	original_language	year	runtime	vote_count	vote_average	IMDB_rating
1	1252		The Dark Knight Rises	en	2012	165	9263	7.6	7.3
	17818		The Avengers	en	2012	143	12000	7.4	7.2
	19735		Life of Pi	en	2012	127	5912	7.2	6.8
	19971		The Hobbit: An Unexpected Journey	en	2012	169	8427	7.0	6.7
	19261		Skyfall	en	2012	143	7718	6.9	6.6
	19071		Brave	en	2012	93	4760	6.7	6.4
	18723		21 Jump Street	en	2012	109	4151	6.7	6.3
	19129		The Amazing Spider-Man	en	2012	136	6734	6.5	6.3
	19394		Looper	en	2012	118	4777	6.6	6.3
	18972		Men in Black 3	en	2012	106	4228	6.3	6.1

The Simple Recommender

Ranking & Filter

PART TWO



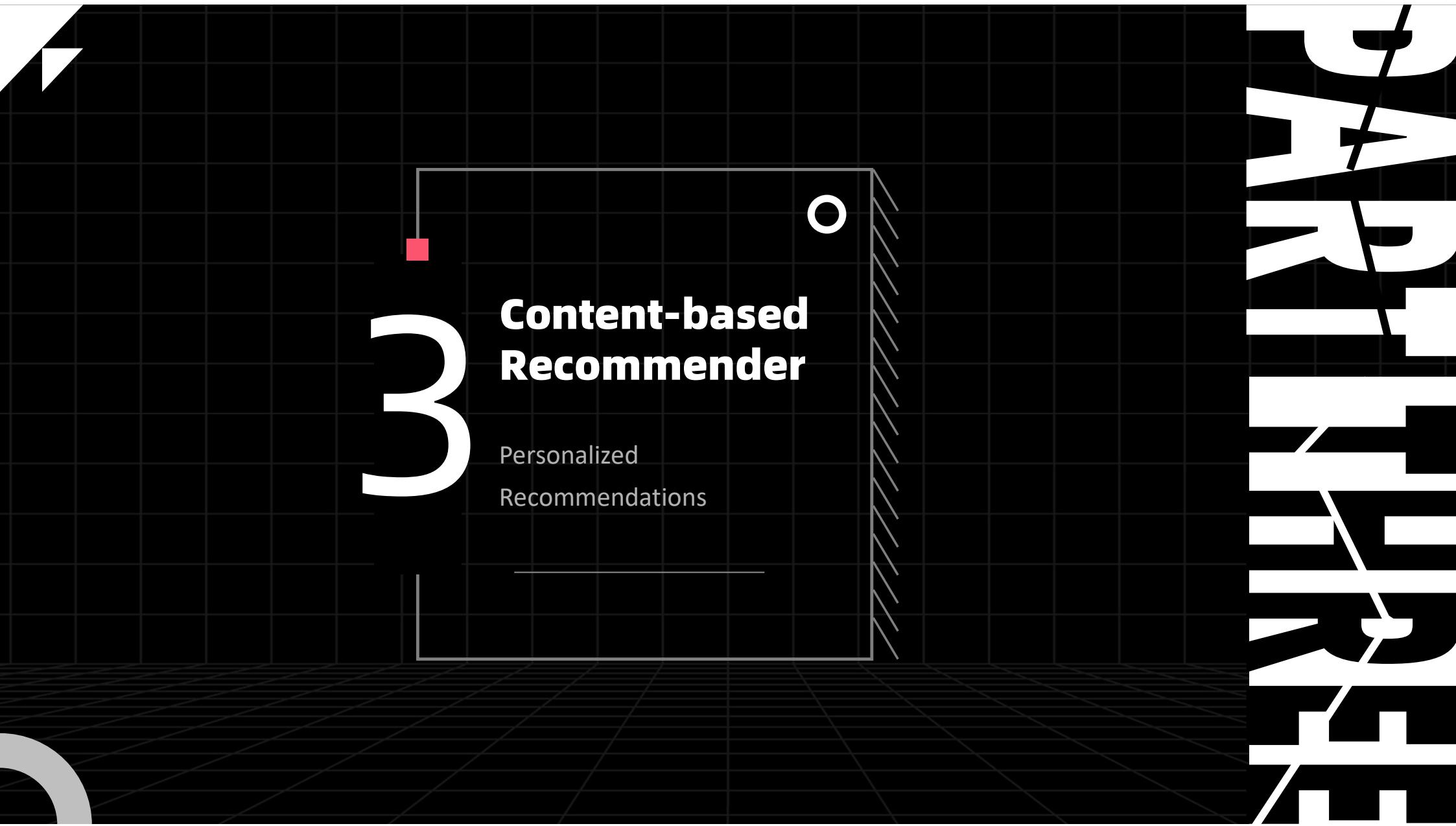
```
build_director_chart("Christopher Nolan", 0.85)
```

		title	original_language	year	runtime	vote_count	vote_average	director	IMDB_rating
12525		The Dark Knight	en	2008	152	12269	8.3	Christopher Nolan	8.3
4110		Memento	en	2000	113	4168	8.1	Christopher Nolan	8.1
1115		Inception	en	2010	148	14075	8.1	Christopher Nolan	8.1
22950		Interstellar	en	2014	169	11187	8.1	Christopher Nolan	8.1
11399		The Prestige	en	2006	130	4510	8.0	Christopher Nolan	8.0
18324		The Dark Knight Rises	en	2012	165	9263	7.6	Christopher Nolan	7.6
10158		Batman Begins	en	2005	140	7511	7.5	Christopher Nolan	7.5
44753		Dunkirk	en	2017	107	2712	7.5	Christopher Nolan	7.4
2474		Following	en	1998	69	363	7.2	Christopher Nolan	6.9
5274		Insomnia	en	2002	118	1181	6.8	Christopher Nolan	6.7
25980		Doodlebug	xx	1997	3	105	6.6	Christopher Nolan	6.2
25981		Doodlebug	xx	1997	3	105	6.6	Christopher Nolan	6.2

3

Content-based Recommender

Personalized
Recommendations



Content-based Recommender

Versions & Calculation Methods

PART THREE



Versions

1. Based on Movie Overviews and Taglines
2. Based on Movie Cast, Crew, Keywords and Genre
3. Based on Watch History

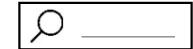
- Cosine Similarity with TF-IDF
Fast and accurate
- Kullback-Leibler divergence
asymmetrical, give more information, but slow
- **Cosine Similarity with CountVectorizer**
Super Fast
- Word2Vec, Doc2Vec
both models generated fast



Content-based Recommender

Versions & Calculation Methods

PART THREE



```
get_recommendations_cs('The Dark Knight') # basic
```

18750	The Shrine
9235	The Hole
28871	Des roses en hiver
21908	Mood Indigo
28720	Vuonna 85
38831	Vinodentro
33337	Let Go
14668	The Princess and the Frog
41438	Our Gang
44257	The Most Hated Woman in America
17730	The Autobiography of Nicolae Ceausescu
26638	Chestnut: Hero of Central Park
3862	Haunted
15667	The Karate Kid
33282	Mary of Nazareth
1957	Flight of the Navigator
25279	Listen Up Philip
42264	Trancers 6: Life After Deth
20855	Big Girls Don't Cry... They Get Even
25859	Fighting Mad
Name: title, dtype: object	

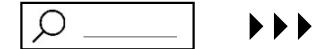
```
get_recommendations_cs('The Dark Knight', recommender = alter) :
```

18442	The Dark Knight Rises
10210	Batman Begins
11463	The Prestige
2486	Following
15651	Inception
4126	Memento
5302	Insomnia
164	Feast of July
20253	Jack Reacher
3799	The Way of the Gun
10241	Retrograde
23668	Mission: Impossible - Rogue Nation
8243	Whispers in the Dark
10639	Creep
463	The Englishman Who Went Up a Hill But Came Dow...
34	Carrington
3814	Best in Show
11348	Severance
2456	Wing Commander
32829	Scouts Guide to the Zombie Apocalypse
Name: title, dtype: object	

Content-based Recommender

Versions & Calculation Methods

PART THREE



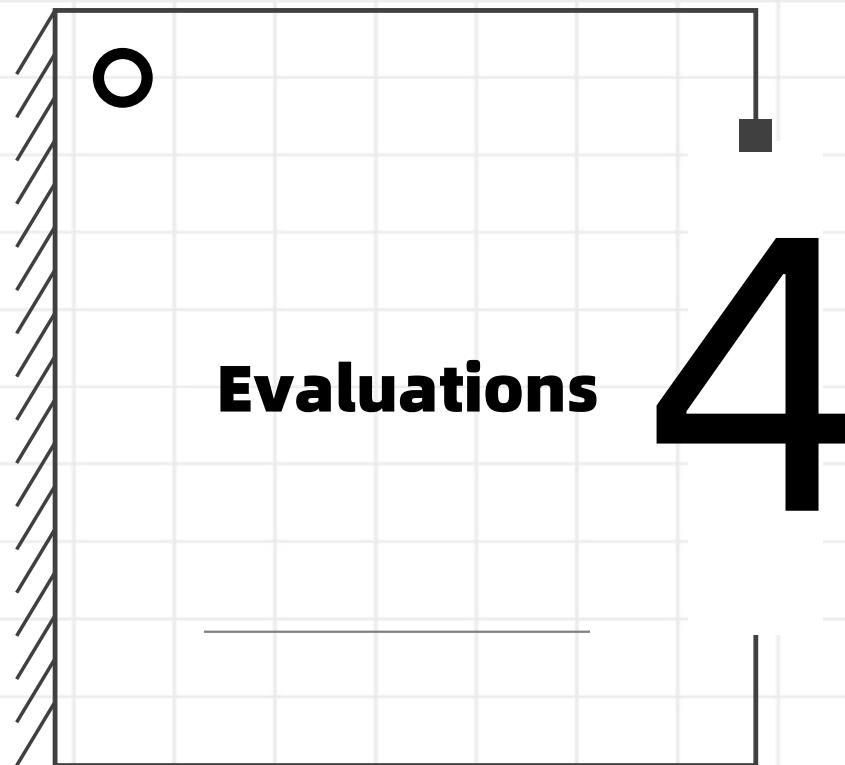
```
get_recommendations_d2v(10530, 5)
```

```
['Heaven', 'Julius Caesar', 'The Simpsons Movie', 'Man of Aran', 'Jane Eyre']
```

```
get_recommendations_d2v(11111, 10)
```

```
['A Countess from Hong Kong',
'César et Rosalie',
'De Lift',
'No Way Out',
'Rembrandt',
'One False Move',
"It's My Party",
'The Black Stallion',
'Into the West',
'Alice']
```

3
2
1
0
9
8
7
6
5
4
3
2
1



Evaluation

Approaches & Results

PART FOUR



Plan

Simple Recommender
Compared with ranking from IMDb website

Content-based Recommender
Compared with recommendations on Amazon Prime

Evaluation

Simple Recommender

PART FOUR



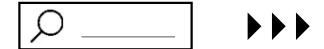
Rank & Title	IMDb Rating
1. The Shawshank Redemption (1994)	★ 9.2
2. The Godfather (1972)	★ 9.2
3. The Dark Knight (2008)	★ 9.0
4. The Godfather: Part II (1974)	★ 9.0
5. 12 Angry Men (1957)	★ 8.9
6. Schindler's List (1993)	★ 8.9

title
314 The Shawshank Redemption
834 The Godfather
12481 The Dark Knight
292 Pulp Fiction
2843 Fight Club
522 Schindler's List
351 Forrest Gump
23673 Whiplash
3030 The Green Mile
289 Leon: The Professional

Evaluation

Content-based Recommender

PART FOUR

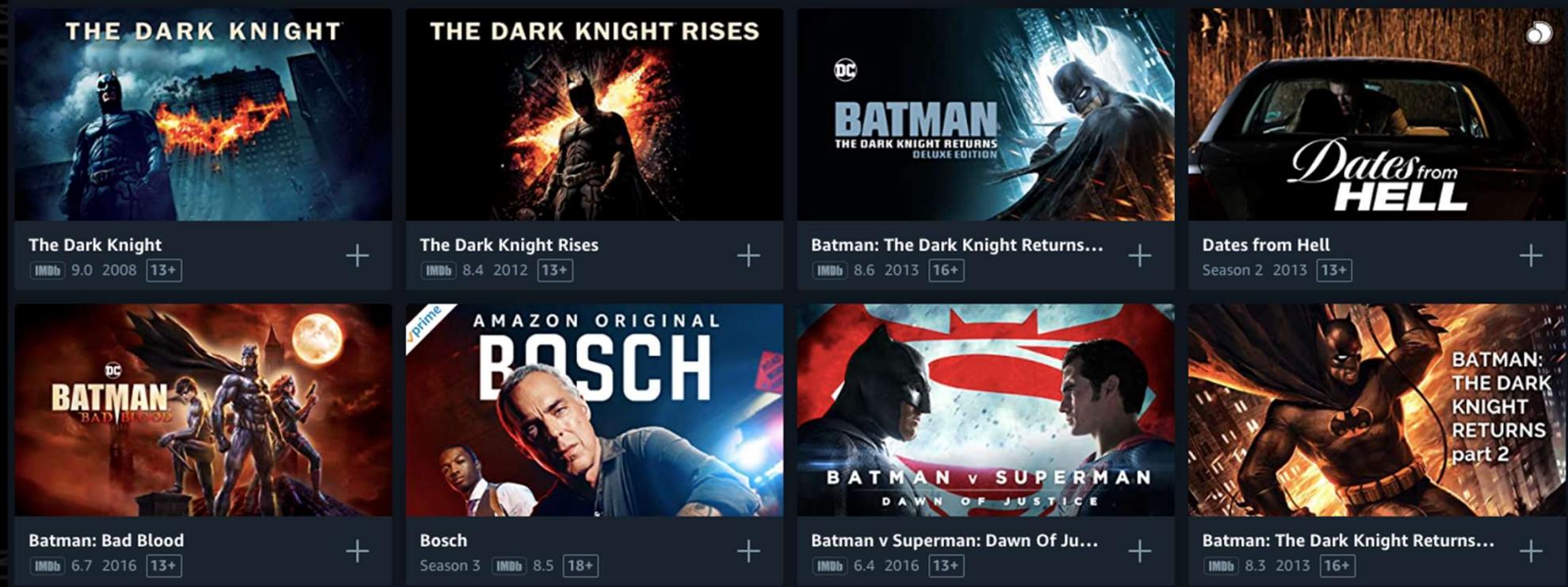


get_recommendations_cs('The Dark Knight') # based on overview and		get_recommendations_cs('The Dark Knight', recommender = alter) # based on	
18750	The Shrine	18442	The Dark Knight Rises
9235	The Hole	10210	Batman Begins
28871	Des roses en hiver	11463	The Prestige
21908	Mood Indigo	2486	Following
28720	Vuonna 85	15651	Inception
38831	Vinodentro	4126	Memento
33337	Let Go	5302	Insomnia
14668	The Princess and the Frog	164	Feast of July
41438	Our Gang	20253	Jack Reacher
44257	The Most Hated Woman in America	3799	The Way of the Gun
17730	The Autobiography of Nicolae Ceausescu	10241	Retrograde
26638	Chestnut: Hero of Central Park	23668	Mission: Impossible - Rogue Nation
3862	Haunted	8243	Whispers in the Dark
15667	The Karate Kid	10639	Creep
33282	Mary of Nazareth	463	The Englishman Who Went Up a Hill But Came Dow...
1957	Flight of the Navigator	34	Carrington
25279	Listen Up Philip	3814	Best in Show
42264	Trancers 6: Life After Deth	11348	Severance
20855	Big Girls Don't Cry... They Get Even	2456	Wing Commander
25859	Fighting Mad	32829	Scouts Guide to the Zombie Apocalypse
Name: title, dtype: object		Name: title, dtype: object	

Evaluation

Content-based Recommender

PART FOUR



Evaluation

Content-based Recommender

PART FOUR



get_recommendations_cs('Jurassic Park', n=20, vote = 7, recommender=cs)		get_recommendations_cs('Jurassic Park', recommender=alter)	
25083	United States of Secrets (Part One): The Program	1494	The Lost World: Jurassic Park
21607	Growing Pains	4276	A.I. Artificial Intelligence
30716	Ghost in the Shell: Stand Alone Complex - The ...	522	Schindler's List
11644	The Weeping Meadow	8918	Twilight Zone: The Movie
2680	Stiff Upper Lips	4477	Always
20990	The Colors of the Mountain	1634	Amistad
36030	Forgotten	18345	War Horse
1033	Secrets & Lies	39904	The BFG
1654	Other Voices Other Rooms	5913	Catch Me If You Can
44208	The Evil Within	5359	Minority Report
3450	Cabaret	18259	The Adventures of Tintin
20506	Stella Maris	1086	E.T. the Extra-Terrestrial
4391	Camille Claudel 1915	1359	Jaws
23062	Highway	7021	1941
41983	Love Finds You In Sugarcreek	3643	The Sugarland Express
4317	The Sons of Katie Elder	12733	Indiana Jones and the Kingdom of the Crystal S...
14844	Swades: We, the People	2651	The Color Purple
9343	Is It Fall Yet?	1267	Indiana Jones and the Last Crusade
41573	Erebus: Operation Overdue	7906	The Terminal
38001	Fimfarum 2	1177	Raiders of the Lost Ark
Name: title, dtype: object		Name: title, dtype: object	

Evaluation

Content-based Recommender

PART FOUR



Jurassic Park

IMDb 8.2 1993 13+



The Lost World: Jurassic Park

IMDb 6.5 1997 13+



Jurassic Park III

IMDb 5.9 2001 13+



Jurassic World

IMDb 6.9 2015 16+



Jurassic World: Fallen Kingdom

IMDb 6.1 2018 13+



Toy Hunter

Season 1 2012 ALL



E.T.,The Extra-Terrestrial

IMDb 7.9 1982 7+



Back to the Future

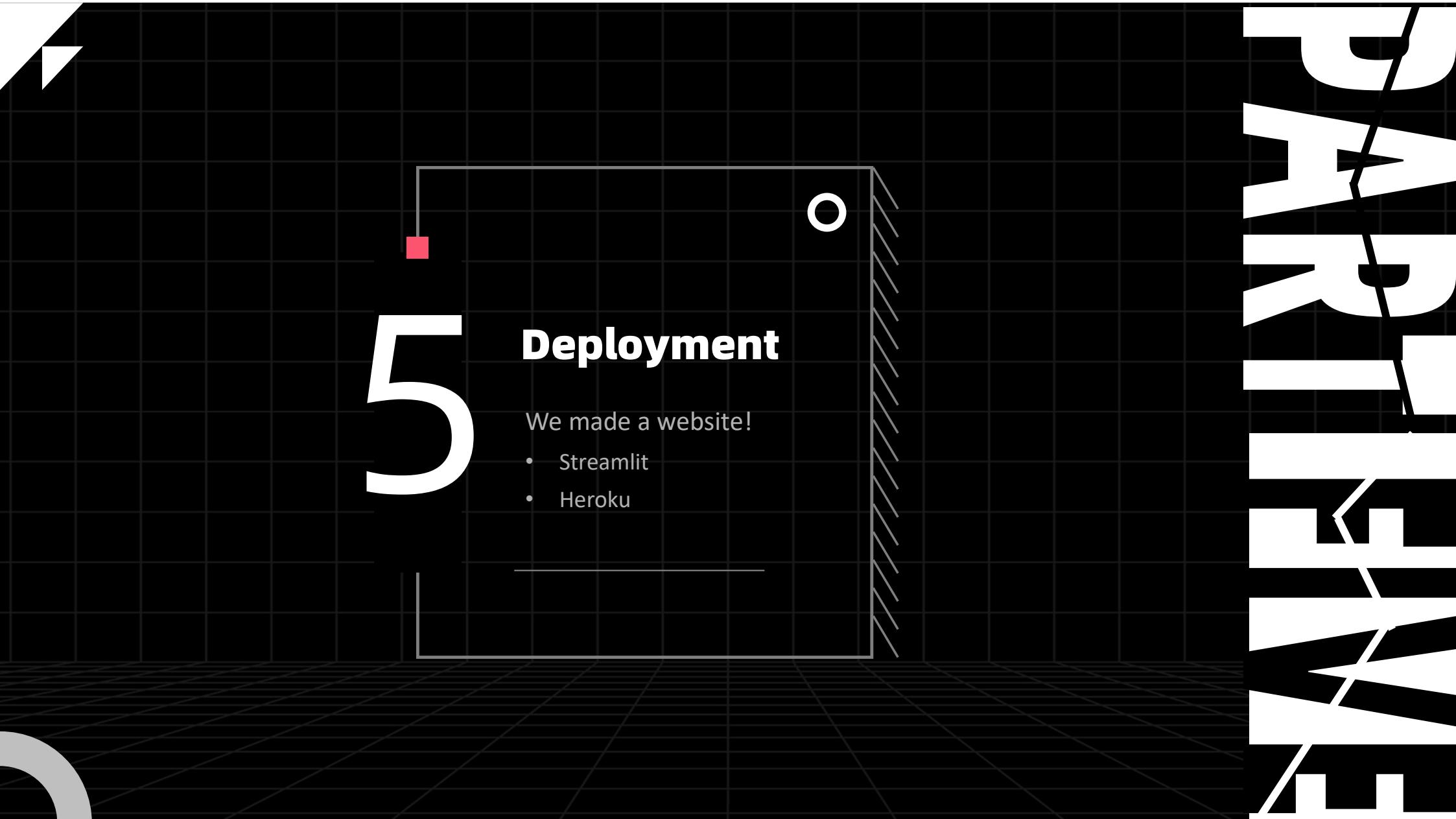
IMDb 8.5 1985 13+

5

Deployment

We made a website!

- Streamlit
- Heroku



Website

PART FIVE



01

Simple Recommender

Simple Movie Recommender

Select your favorite genre

Action

Set your preferred movie runtime in minutes

90

60 240

Recommend!

The Lord of the Rings: The Return of the King | The Lord of the Rings: The Two Towers | The Lord of the Rings: The Fellowship of the Ring | Raiders of the Lost Ark | Pirates of the Caribbean: Dead Man's Chest



02

Content-based Recommender

Movie Recommender System

Type or select a movie from the dropdown

Toy Story ▾

Insert the the minimum score of the movies

0,00

-

+

Show Recommendation

Future Plan

- More Comparisons
- Better Evaluation
- Combine both
recommenders

6

Thank You For Listening

Recommender

• • •

Movie

