

# Cult Film Recommender

...

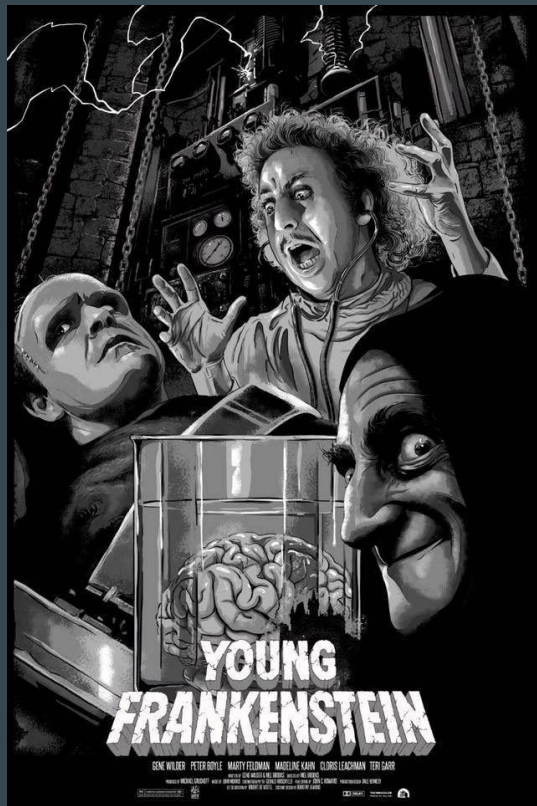
Emily Hillenbrand

# Cult Film Recommender

# What makes a Cult Film?

- Has a cult following
- Not an initial commercial success
- Featured on Mystery Science Theater 3000\*

# What is a Cult Film?



# Cult Film Recommender

# Recommendation Systems

Content-Based Filtering

Collaborative Filtering

# Recommendation Systems

## Content-Based Filtering

1. Create system for describing items based on features
2. Manually collect features on items
3. Recommend more of the same

## Collaborative Filtering

# Recommendation Systems

## Content-Based Filtering

1. Create system for describing items based on features
2. Manually collect features on items
3. Recommend more of the same

## Collaborative Filtering

- Find other people who like the same stuff as you
  - Requires a huge user base to get good results
  - Need to submit lots of likes and dislikes



# A New Way to Recommend

How It Works

1. Scrape movie info from reviews
2. Use NLP to figure out the relevant features
3. Map distribution of features for each movie

# A New Way to Recommend

What It Is

- Finds the most similar but with added interface that allows direct feedback
- Created from IMDb user base



HOW I DID IT  
By  
Victor Frankenstein

# Pipeline

Collect Data

Topic Modeling

Output

# Pipeline

## Scrape the Web

- Wikipedia: list of cult films
- IMDb: user reviews

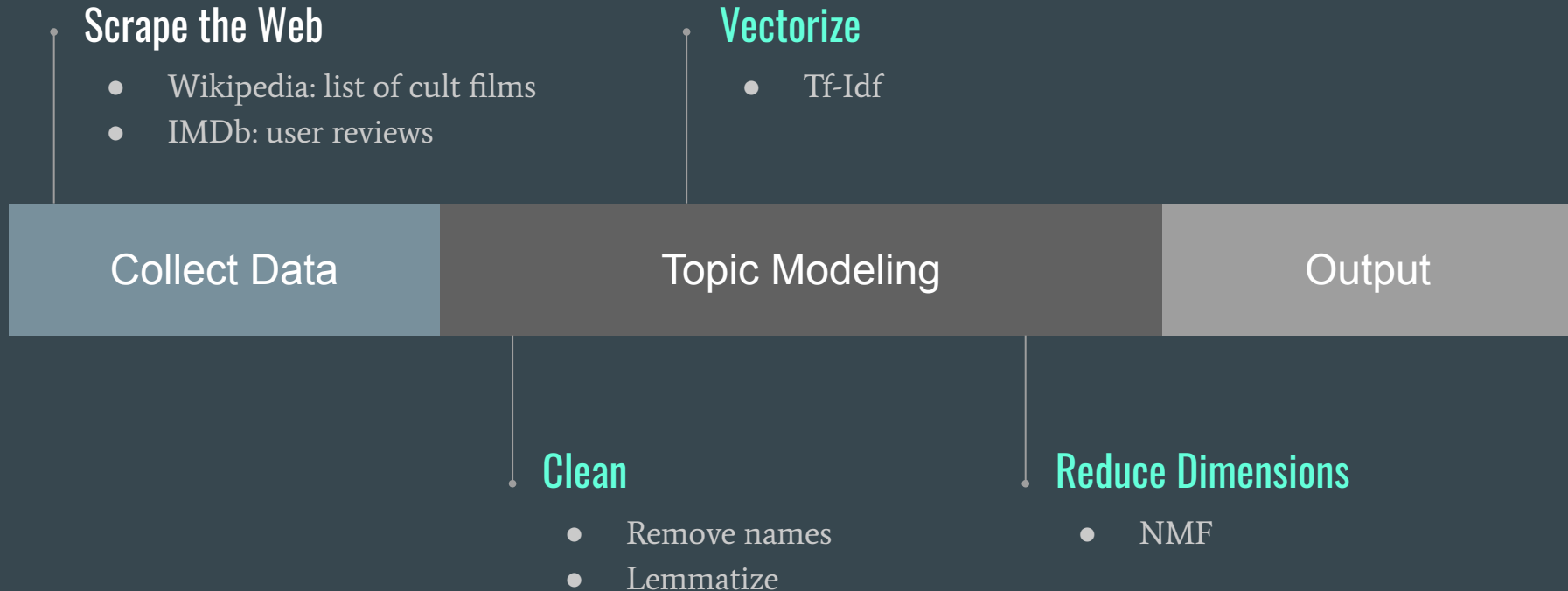
Collect Data

Topic Modeling

Output

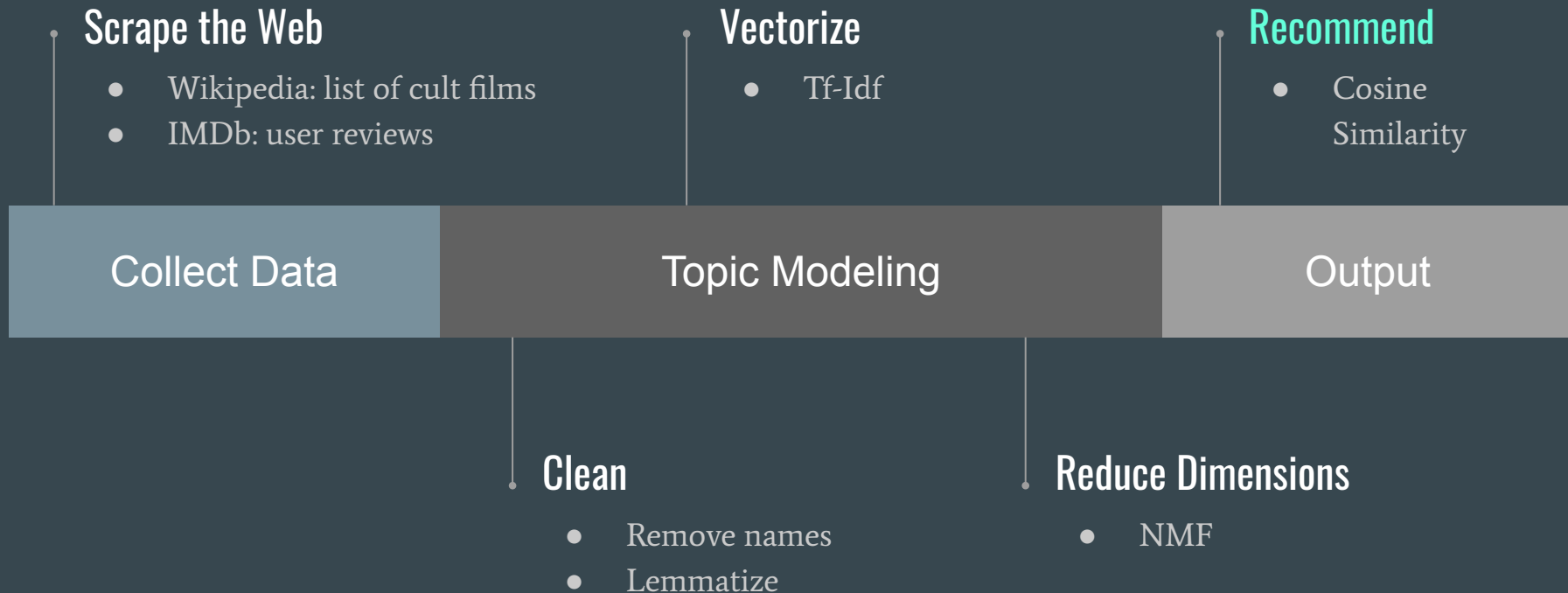
Tools: python, BeautifulSoup, Selenium

# Pipeline



Tools: spaCy, scikit-learn, pandas

# Pipeline



Tools: scikit-learn, jupyter widgets

Topic: ' Art/Classic '

french, art, cinema, war, image

Topic: ' Gory/Scary '

horror, gore, slasher, halloween, house

Topic: ' Sci Fi '

alien, sci\_fi, planet, space, science

Topic: ' Vampires '

vampire, dracula, horror, hammer, blood

Topic: ' Silly '

comedy, funny, laugh, joke, humor

Topic: ' Animated '

animation, animate, anime, disney, cartoon

Topic: ' Martial Arts '

martial, kung\_fu, fight, action, kong

Topic: ' Music  
Documentary '

band, music, concert, rock, song

Topic: ' Living Dead '

zombie, horror, gore, fulci, shaun

Topic: ' Giant Monster '

monster, godzilla, creature, frankenstein,  
giant

Topic: ' Russia '

russian, soviet, polish, russia, poland

Topic: ' Sex '

sex, woman, porn, gay, sexual

Topic: ' Gold '

leprechaun, lep, gold, cody, vegas

Topic: ' Action '

action, cop, gang, western, noir

Topic: ' Kids '

school, teen, kid, girl, high

Topic: ' Holiday '

christmas, santa, claus, holiday, martian





# Example Movie: Shaun of the Dead

“A Romantic Comedy With Zombies”

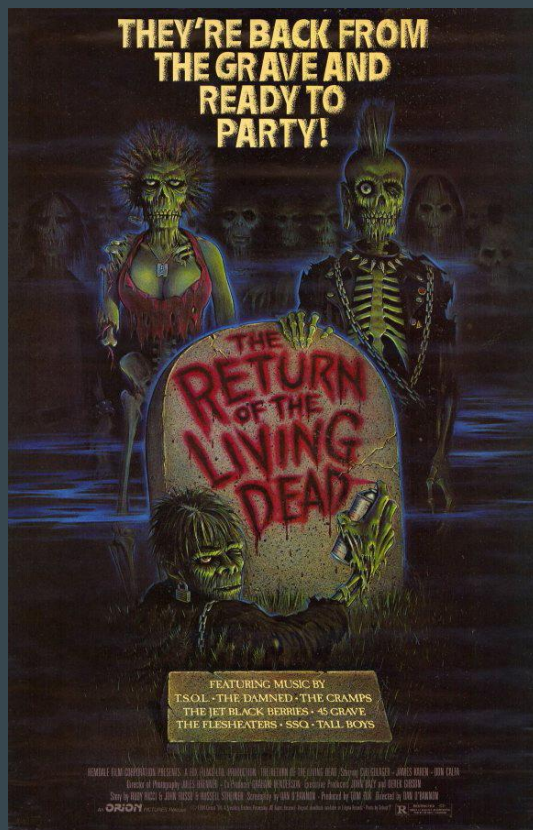
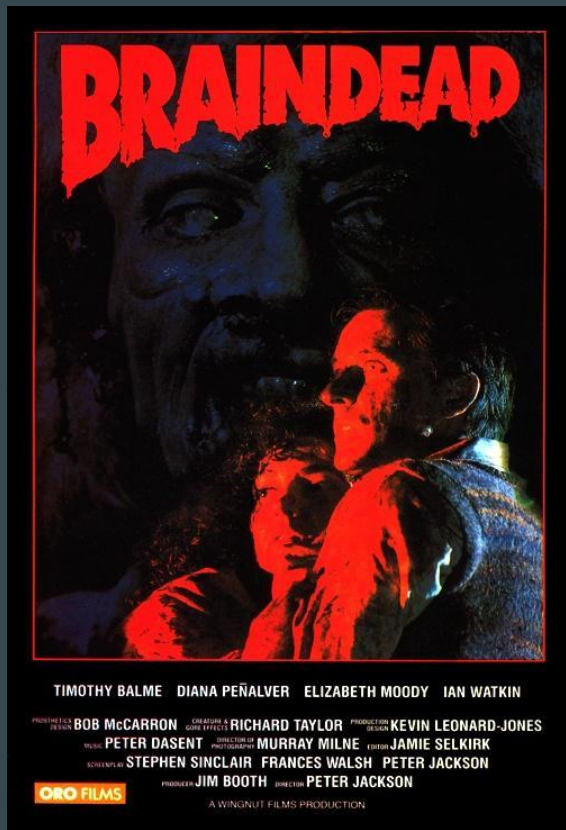
- British zombie horror comedy
- Balances silly with serious



Sci Fi	<input type="text" value="1.25"/>	1.25	1.2514
Vampires	<input type="text" value="0.47"/>	0.47	0.4670
Silly	<input type="text" value="41.59"/>	41.59	41.586
Animated	<input type="text" value="0.00"/>	0.00	0
Martial Arts	<input type="text" value="0.00"/>	0.00	0
Music Documentary	<input type="text" value="0.00"/>	0.00	0
Living Dead	<input type="text" value="90.92"/>	90.92	90.924
Giant Monster	<input type="text" value="0.00"/>	0.00	0
Russia	<input type="text" value="0.00"/>	0.00	0
Sex	<input type="text" value="0.00"/>	0.00	0
Gold	<input type="text" value="0.09"/>	0.09	0.0938
Action	<input type="text" value="0.00"/>	0.00	0
Kids	<input type="text" value="0.00"/>	0.00	0
Holiday	<input type="text" value="0.00"/>	0.00	0

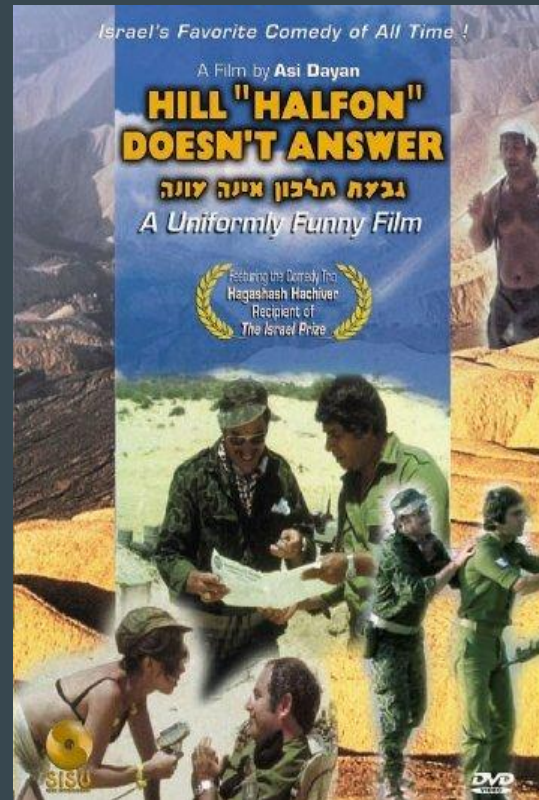
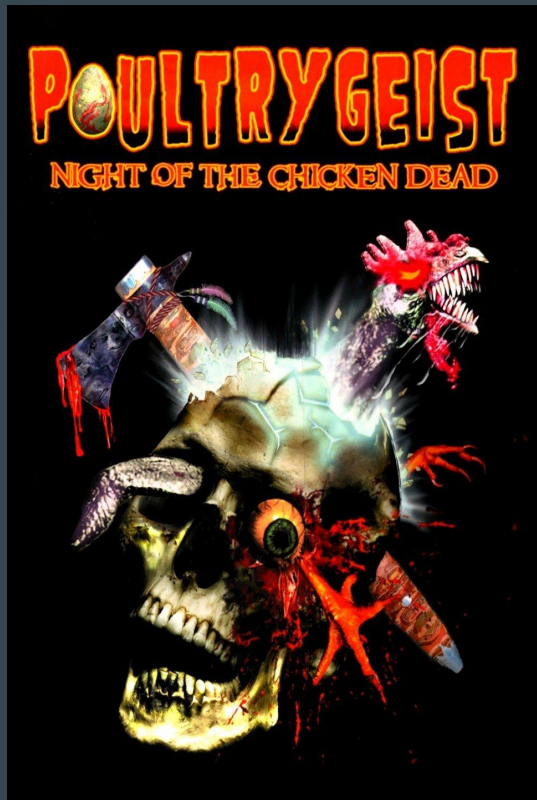
1 ▶ # get recommended movies↕

executed in 157ms, finished 01:13:35 2019-12-10



Most similar to Shaun of the Dead





Much sillier than Shaun of the Dead

# Other Applications

Creating Features from Reviews

- Amazon Products
- Yelp Restaurants
- Comic Books
- Anything else with reviews

# Thank You

What are your favorite cult films?

Emily Hillenbrand



/in/hillenbrandemily



/ehillenbrand4541

# Pipeline

## Scrape the Web

- Wikipedia: list of cult films
- IMDb: user reviews

Collect Data

## Vectorize

- Tf-Idf

Topic Modeling

## Recommend

- Cosine Similarity

Output

## Clean

- Remove names
- spaCy Lemmatization

## Reduce Dimensions

- NMF