

MOVIE REVIEW FILTERS

“We want to build trust between the user and their personalized recommendations by enabling them to choose between hundreds of recommendations within seconds.”

This video is the capstone project report created by the participants of an *aggregate intellect* premium hands-on workshop. See <https://ai.science> for more details

2020

Team



Data Scientist
Developed the strategy and
built the recommender
engine.
www.jiristodulka.com



Data Scientist
Developed the LDA model
to extract filters and
performed app deployment.

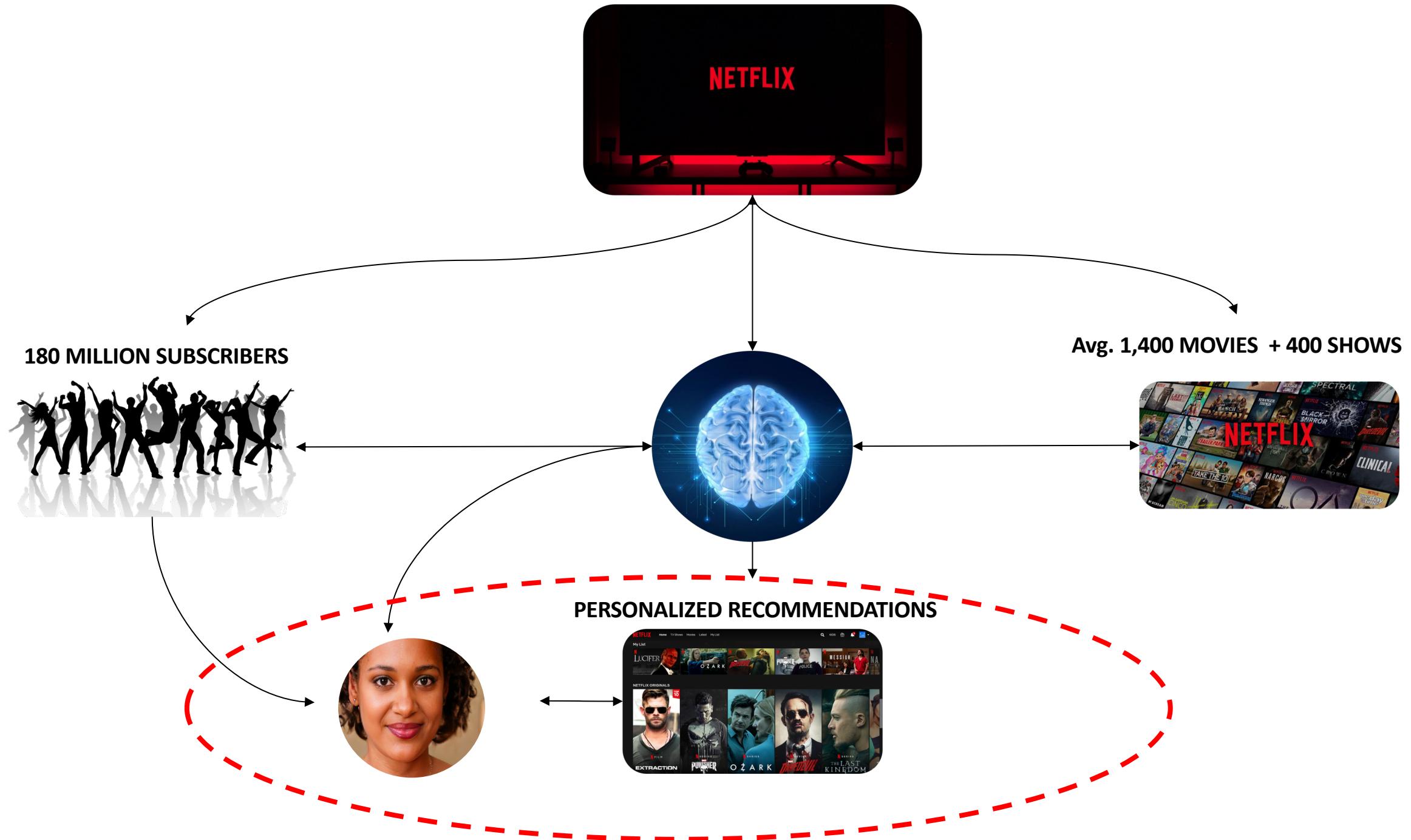


“In linguistics, an adjective is a word that modifies a noun or noun phrase or describes its referent. Its semantic role is to change information given by the noun.”

Wikipedia

“Neuro-imagery scans show that when consumers evaluate brands, they use their emotions, based on personal feeling and experiences, rather than information about the brand, such as features and facts.”

Research

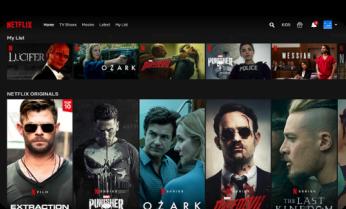


A close-up, slightly blurred portrait of a woman with dark skin, curly hair, and a gentle smile. She is wearing small hoop earrings. The background is dark and out of focus.

Harriet Gan

- Average Netflix user
- Watching time: 2h/day
- 18 minutes: 50 titles

18 minutes in 50 intersections



50 interactions

Reference



Top Picks for Joshua



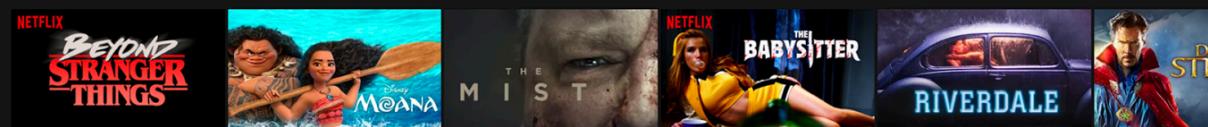
Trending Now



Because you watched Narcos



New Releases



In a cooking showdown, sabotage is the spice. Let's see what the competition can do -- with a vending machine.

+ MY LIST



EPISODES

- Campy Movies
- Children & Family Movies
- Chinese Movies
- Classic Action & Adventure
- Classic Comedies
- Classic Dramas
- Classic Foreign Movies
- Classic Movies
- Classic Musicals
- Classic Romantic Movies
- Classic Sci-Fi & Fantasy
- Classic Thrillers
- Classic TV Shows

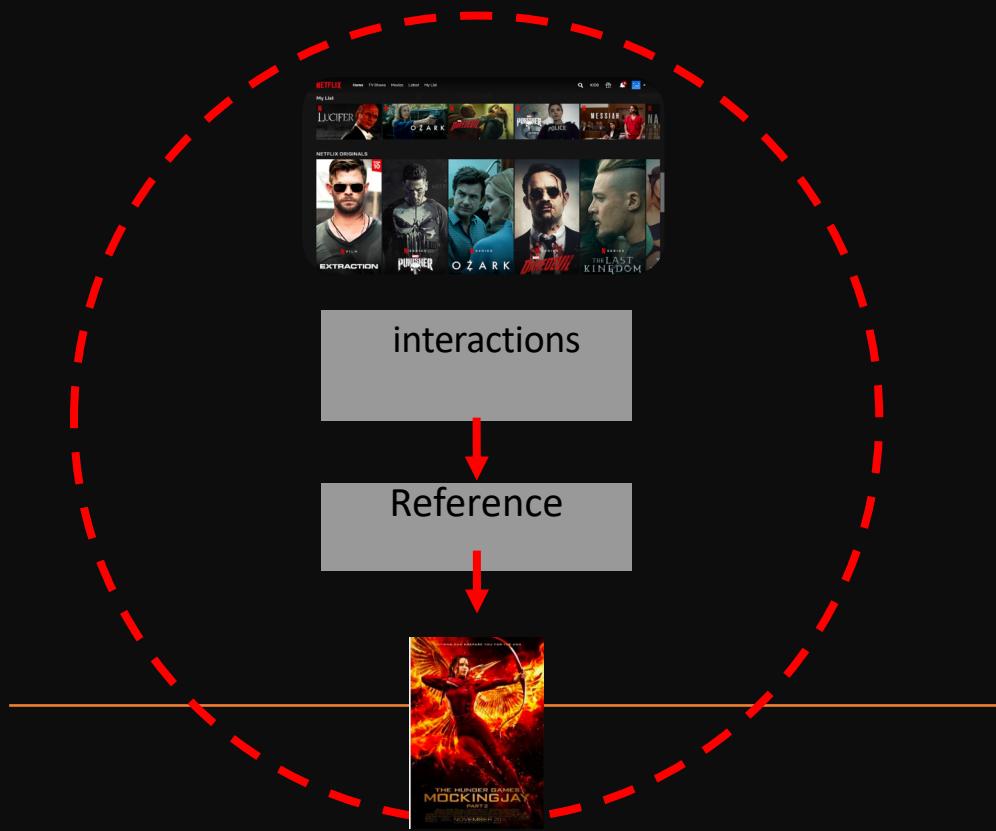
Continue Watch



Harriet wants to invest 2h of her time

External Reference

18 minutes in 50 Interactions



Google

hunger games reviews

- hunger games reviews
- hunger games reviews book
- hunger games reviews movie
- hunger games reviews for parents
- hunger games reviews imdb
- hunger games recenze
- hunger games recenze knihy
- hunger games recenze filmu
- hunger games movies reviews

IMDb

Movies, TV & Showtimes

Celebs, Events & Photos

News & Community

Watchlist

The Hunger Games: Mockingjay - Part 2 (2015)

12A | 137 min | Adventure, Sci-Fi | 19 November 2015 (UK)

Your rating: ★★★★★★★★★★★ /10

7.1 Ratings: 7.1/10 from 38,869 users Metascore: 65/100

Reviews: 178 user | 286 critic | 4 from Metacritic.com

As the war of Panem escalates to the destruction of other districts by the Capitol, Katniss Everdeen, the reluctant leader of the rebellion, must bring together an army against President Snow, while all she holds dear hangs in the balance.

Director: Francis Lawrence

Writers: Peter Craig (screenplay), Danny Strong

Posted by u/Icyrow 2 years ago

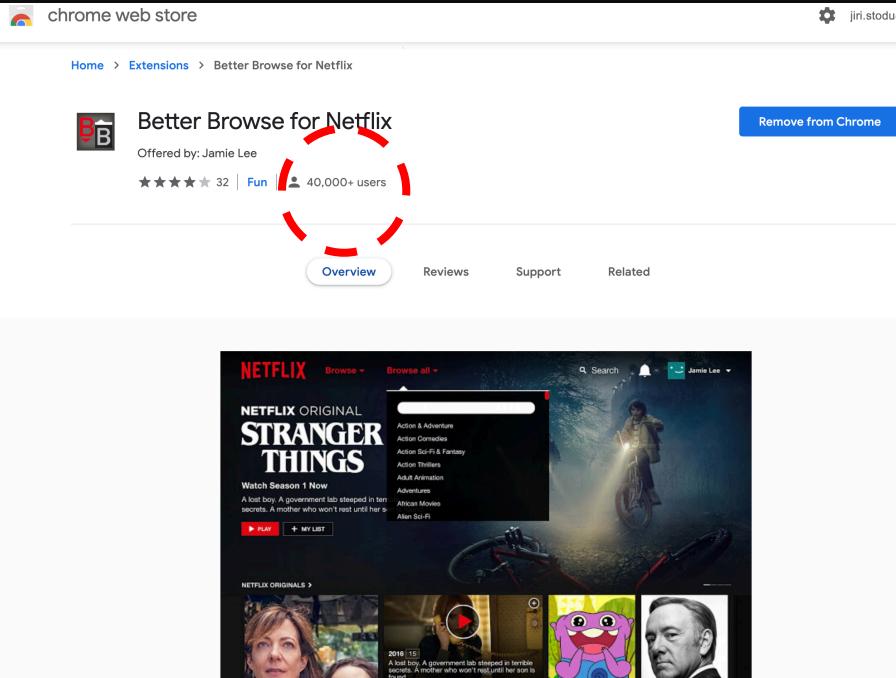
39.7k

Why doesn't netflix have a decent way to browse content? I feel like i'm fairly stuck with the 50-100 titles shown to me on the homescreen, why can't I browse their thousands of titles that they do they have outside of a search bar? why do I have to know the shows name to find it?



2.1k Comments Give Award Share Save Hide Report

94% Upvoted



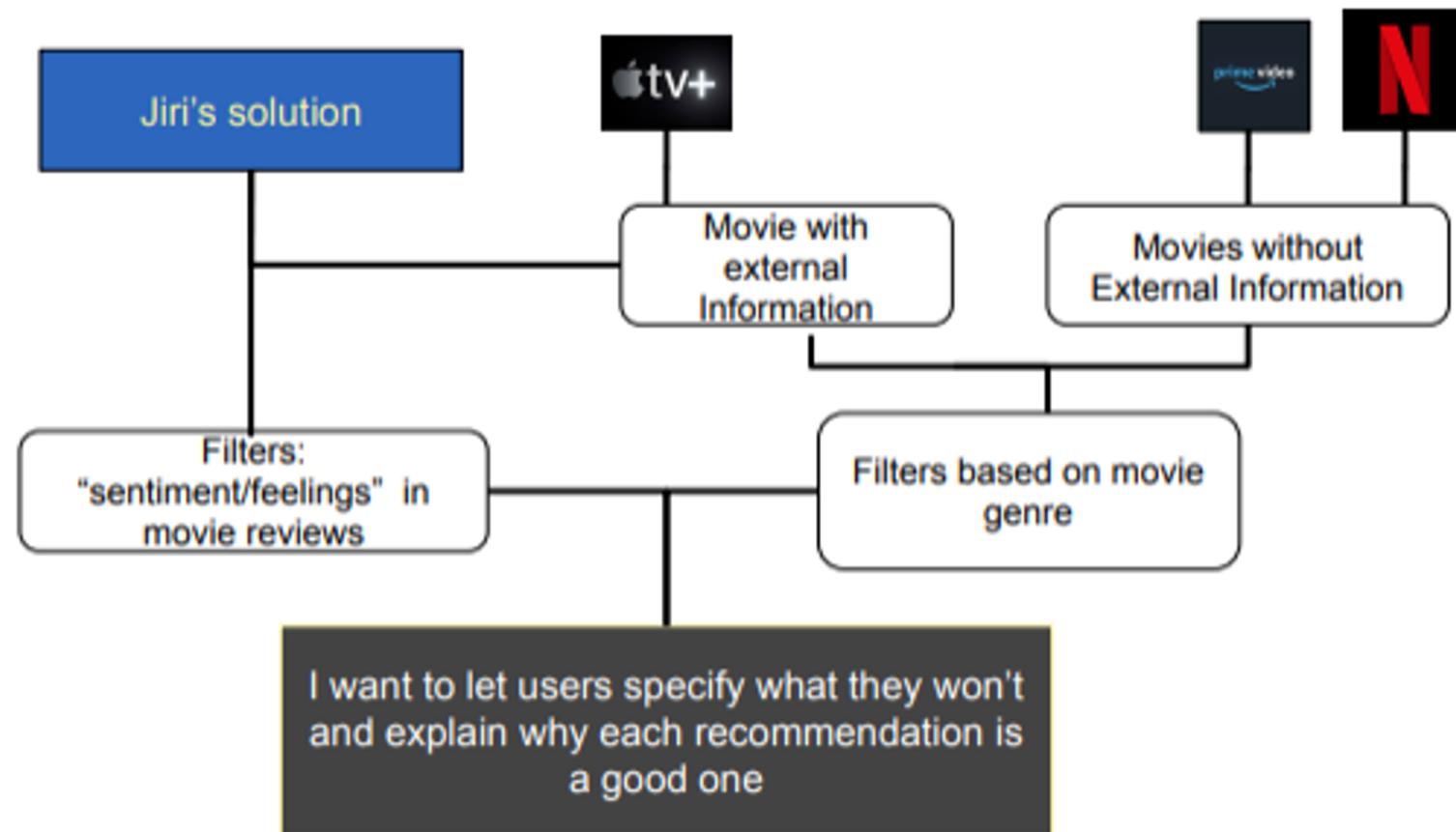
MARKET NEEDED

- 180 mill Netflix users worldwide
- Early adopters: 1,000s -> web app
- Niche market: 10,000s -> Chrome Extension

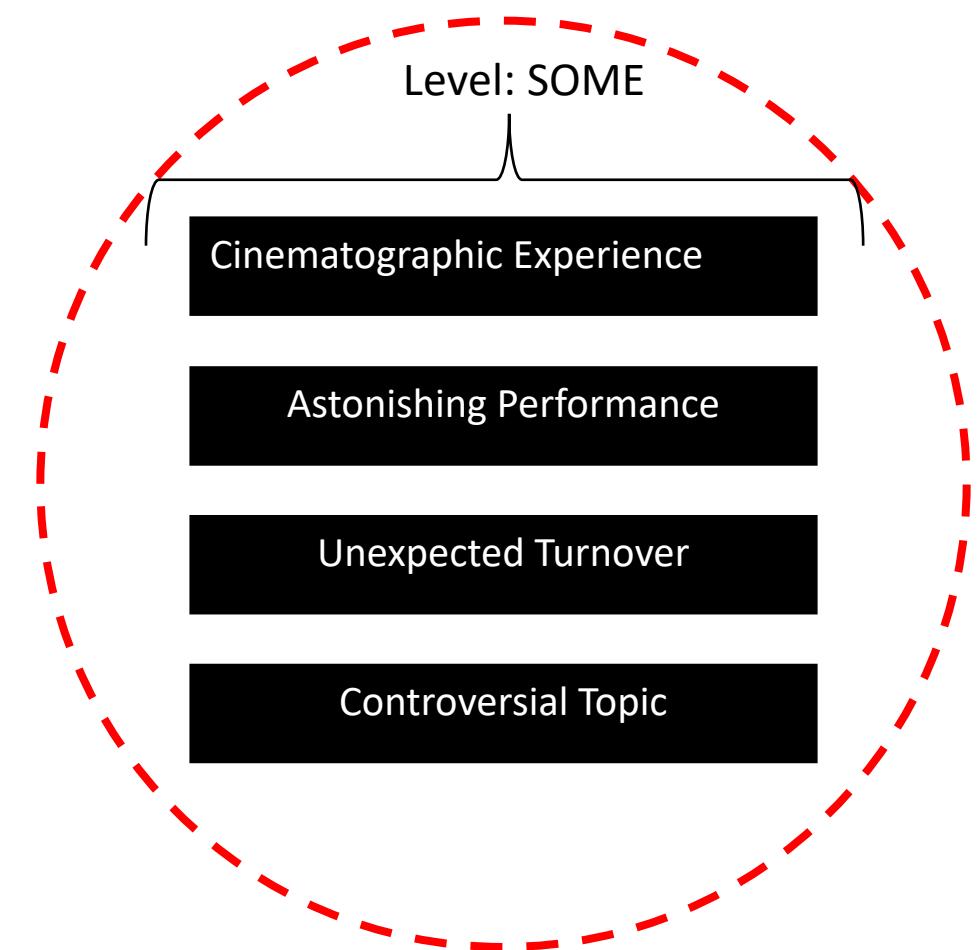
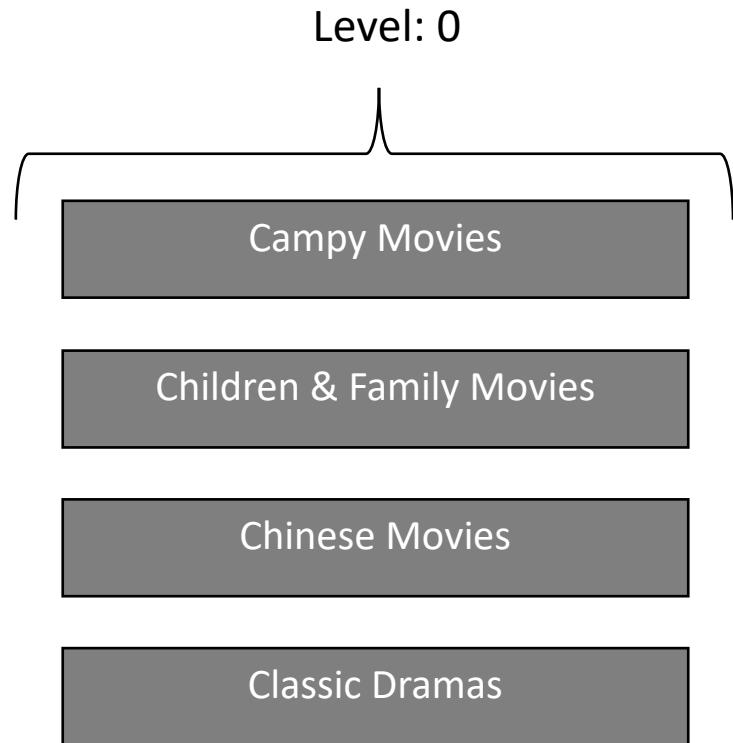
Market Need

- **Early Adopters:**
 - **Film Aficionado in their 20s and 30s. Impulsive in their movie habits.**
 - **Want to spend less time for enjoyment (better movies).**
 - **Often visit external websites for reviews and ratings.**
- **Broader market: 180 million Netflix Users**
 - **Much less picky.**
 - **But still wants a one click solution if it enhances their experience.**

Idea Maze



Browse via Emotional Experience



Can We Help Harriet Even More?

Cinematographic Experience

4

Astonishing Performance

4

Unexpected Turnover

5

Controversial Topic

0



0 5



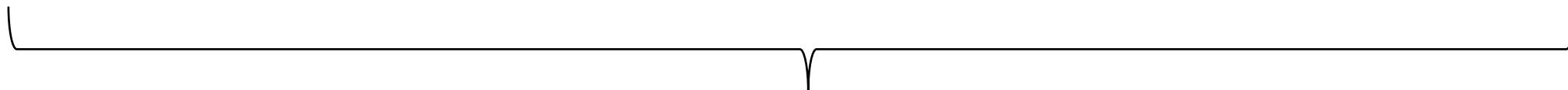
0 5



0 5



0 5



Emotional Level: QUANTIFIED!!!!

Harriet can Invest and Benefit More

- <https://clockify.me/time-tracking-statistics>

Entire Target Customers: Invest and Benefit More

- <https://clockify.me/time-tracking-statistics>

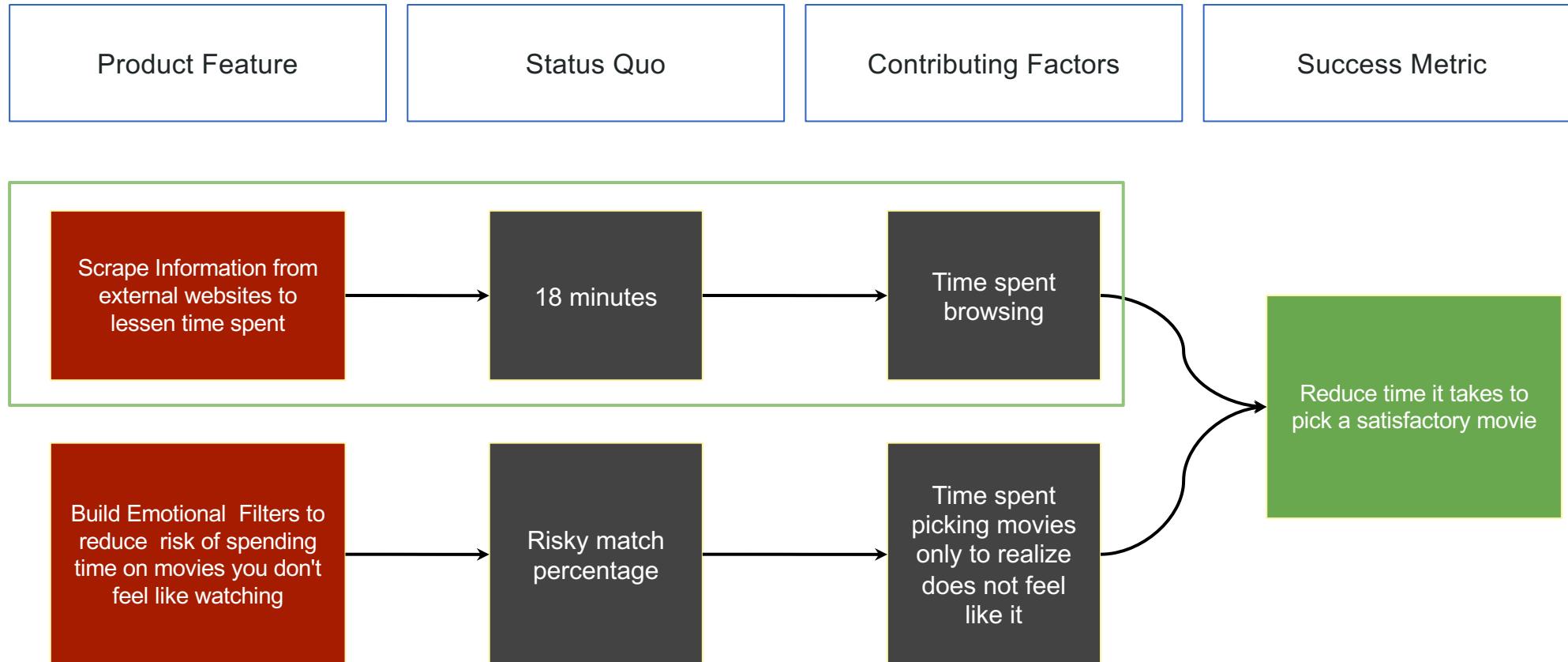
Business

- Open source app/Chrome Extension
- Product for Netflix (IP)

Product

- Call for the session based dataset
- 18 min before final choice
- Treatment vs. control of people
- Try the app before us

Product Strategy



Cards at Hand

- Accessible Data
- Data Wrangling
- Scraping Data
- Desired Data

IMDb > Argo (2012) > Reviews & Ratings - IMDb

Reviews & Ratings for
Argo More at IMDbPro »

Filter: Prolific Authors Hide Spoilers:

Interleaved...
Reviews from users who have written at least 100 reviews, most prolific authors first.
Reviews from users who have written at least 100 reviews, most prolific authors first.

Page 1 of 23: [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] ▶

Index 222 matching reviews (656 reviews in total)

2 out of 12 people found the following review useful:

 **Oddly, I couldn't find anything to dislike about this film!**

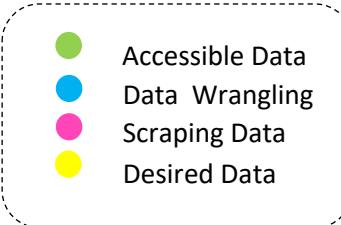
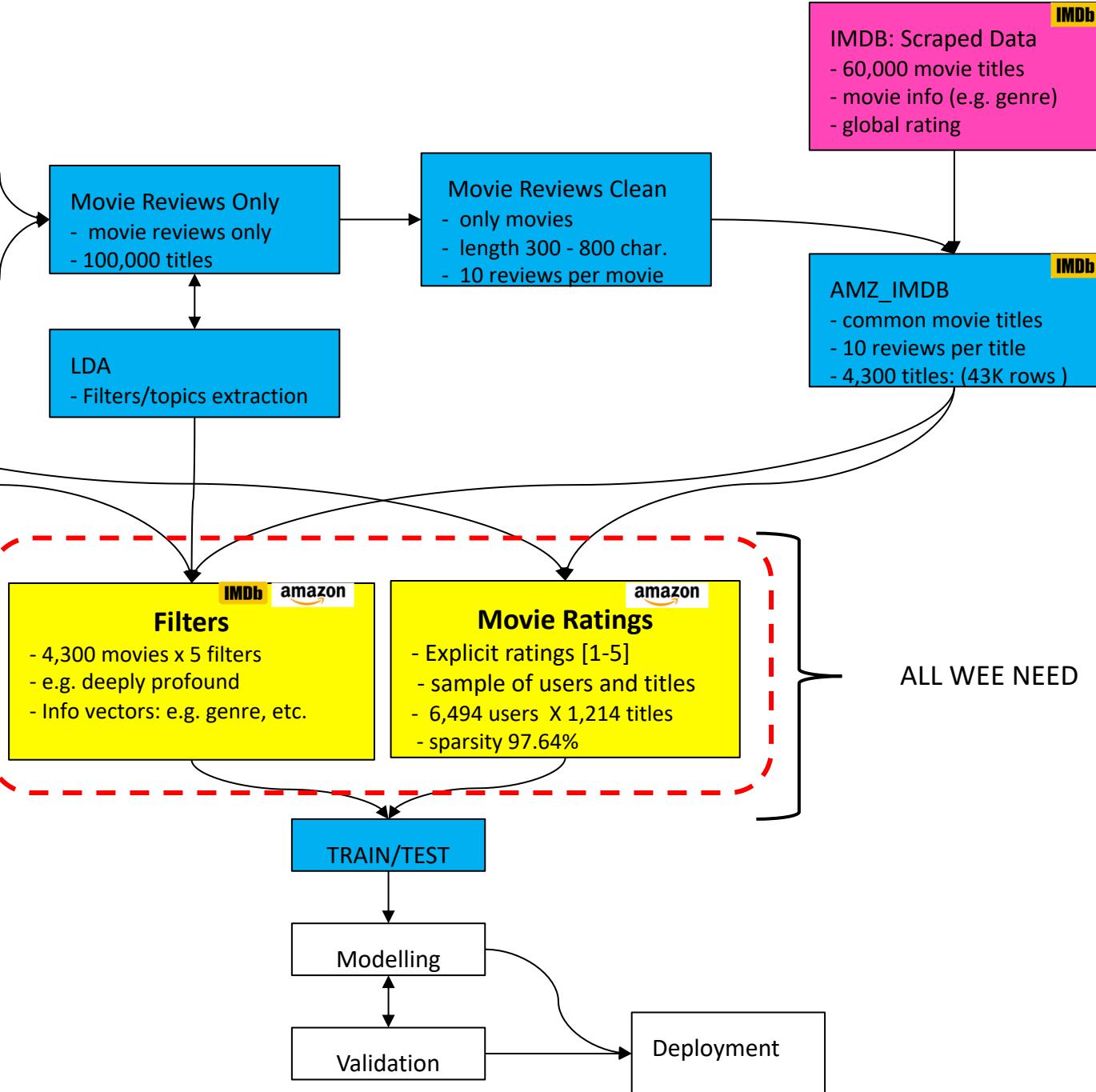
Author: planktonrules from Bradenton, Florida
18 February 2013

I will readily admit that I am a very critical person when it comes to movies. After all, a norm counting) reviews to IMDb! However, "Argo" is an unusual film because I honestly can't think Really...it's THAT good! The film is about a joint effort by the Canadian government and the C Iran during their revolution in 1979-80. It seems that most of the Americans in the US embas of folks escaped and sought shelter in the Canadian embassy. What happened next? See t

It's odd. In light of the film's greatest strength, how could the Oscar folks NOT have nominat many ways the film was wonderfully directed. Although I know the fate of the six refugees, I

Filters

Movie Ratings



Filters

- 4,300 movies x 5 filters
- e.g. deeply profound
- Info vectors: e.g. genre, etc.

	Cinematographic Experience	Perfectl fun	Cheesy fun	Experience excitement	Diverse
Guardians of the Galaxy	5	5	2	0	0
The Big Lebowski	0	3	0	1	2
John Wick	3	2	2	3	1
Skyfall	2	0	0	5	2
.
.
.
Mad Max: Fury Road	5	3	0	0	1

Step 1

Mapping the look up words in the reviews

	Filter 1: deep, profound Look up words in reviews: ["deep", "memorable", "happy", "poetic", "noble"]	Filter 2: entertaining music Look up words in reviews: ["entertaining", "musical", "music", "score"]	Filter 3: "realistic settings Look up words in reviews: ["believable", "accurate", "correct", "likable"]	Filter 4: Experience excitement Look up words in reviews: ["special", "worth", "sexual", "mysterious", "dangerous"]	Filter 5: Fun Look up words in reviews: ["dumb", "comic", "funny", "humorous", "hilarious", "clever", "witty"]
review 1: Guardians of the Galaxy	0	1	0	1	0
review 2: Guardians of the Galaxy	0	1	0	1	0
review 3: Guardians of the Galaxy	1	0			1
...
review 8: Guardians of the Galaxy	0	1	0	0	0
review 9: Guardians of the Galaxy	1	1	1	1	0
review 10: Guardians of the Galaxy	1	1	1	0	0

Step 2

Vectorize

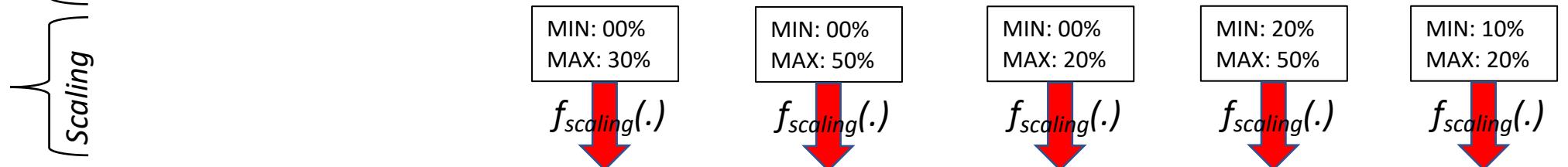
Guardians of the Galaxy	3/10 = 30%	5/10 = 50%	2/10 = 20%	3/10 = 30%	1/10 = 10%
-------------------------	------------	------------	------------	------------	------------

Step 2

Vectorize all movie titles

	Cinematographic Experience	Perfect fun	Cheesy fun	Experience excitement	Diverse
Guardians of the Galaxy	$3/10 = 30\%$	$5/10 = 50\%$	$2/10 = 20\%$	$3/10 = 30\%$	$1/10 = 10\%$
The Big Lebowski	$0/10 = 00\%$	$4/10 = 40\%$	$1/10 = 10\%$	$2/10 = 20\%$	$2/10 = 20\%$
John Wick	$3/10 = 30\%$	$3/10 = 30\%$	$2/10 = 20\%$	$4/10 = 40\%$	$1/10 = 10\%$
Skyfall	$2/10 = 20\%$	$0/10 = 00\%$	$0/10 = 00\%$	$5/10 = 50\%$	$2/10 = 20\%$
.
.
.
Mad Max: Fury Road	$3/10 = 30\%$	$4/10 = 40\%$	$1/10 = 10\%$	$5/10 = 50\%$	$1/10 = 10\%$

Step 3:

*Filters on [0 - 5] scale*

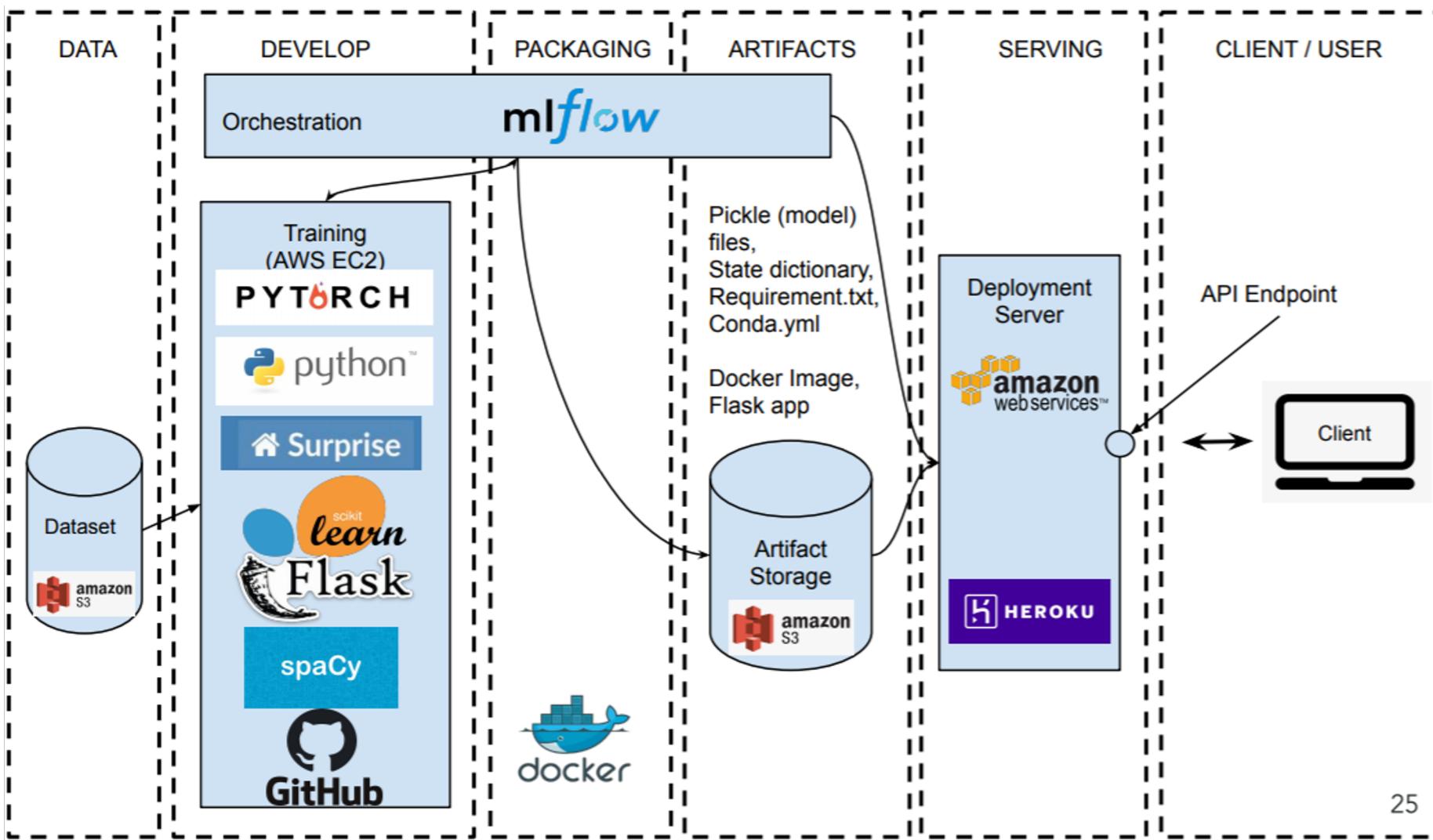
Filters

	Cinematographic Experience	Perfect fun	Cheesy fun	Experience excitement	Diverse
Guardians of the Galaxy	5	5	2	0	0
The Big Lebowski	0	3	0	1	2
John Wick	3	2	2	3	1
Skyfall	2	0	0	5	2
.
.
.
Mad Max: Fury Road	5	3	0	0	1

Baseline Model Performance

- How can we measure our success metric
- Call for the dataset
- MAP@k and MRR@k
- MSE 0.5
- Matrix Factorization vs ML problem: e.g. classification task or autoencoder

Architecture



Architecture Issues

- We planned to but didn't have time to translate our code to PyTorch and package with MLFlow since our model comes from sklearn (sklearn-surprise, KNN).
- We used a Flask App to wrap our application in. This helps us deploy a useful friendly way for users to validate the utility of the app.
- We plan to build a Docker image and container and Push it to Amazon Cloud Instance. This way we can reach Alpha users more readily instead of having to demo it in our local machine. However, an issue we encountered is the amount of cloud space we will need for the prediction ratings (1GB).

Alpha Release

*SHOW DEPLOYMENT ON LOCAL MACHINE

sample user 1 = "Kindle User"

sample user 2 = "mirasreviews"

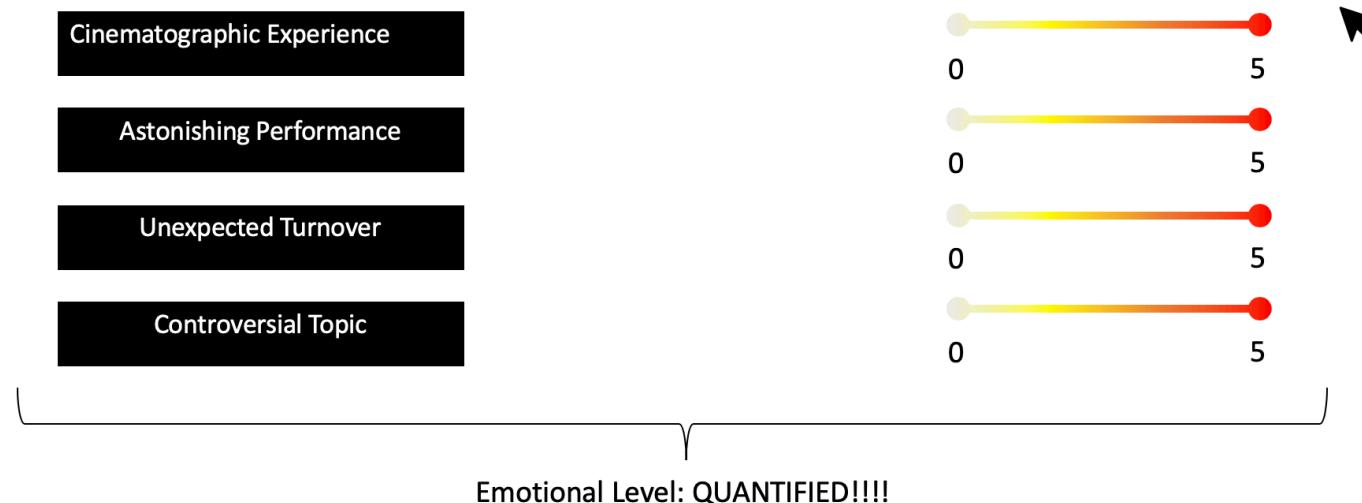
- We are hoping to get feedback from users in terms of how the product was useful/ Or not useful to them considering their username.
- Most importantly, we want to know much much time will a product like this will save the user.

Week 4

Learnings from Alpha Release

Two Users:

- a.) **IDEAL** – “Consumes whatever is likely to satisfy him”
- b.) **PICKY** - “Is very picky and invest into research before consumption ”



Learnings from Alpha Release: Initial Interview

IDEAL

Profile:

- 23 year: 4 years Netflix subscriber
- Weekly time (h): 10

Behaviour:

- Time browsing/searching (min): 0
- Chose top in recommendations
- Never uses IMDB
- Use filters as the last resort

PICKY

Profile

- 25 years: 3 years Netflix subscriber
- Weekly time (h): 13

Behavior:

- Time browsing/searching (min): 10
- Mostly uses external reference (friends)
- Refers to IMDB, FB, Reddit
- Use filters a lot and "doesn't like the recommendations'

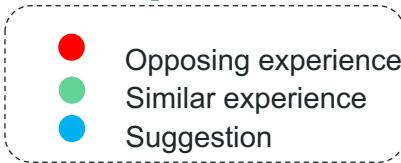
Learnings from Alpha Release: Experience with the App

Questions

- How useful do you think something like this could be (1-10)? Why?
- Do you think this product is user-friendly? Why or why not?
- Do you think this product has good UI? Why or why not?
- After using this product, will you be comfortable with making a movie decision without visiting external app?
- What other filters do you think might be relevant to you?

IDEAL

- 1, **does not understand the words**
- NO, would expect better (coherent) experience
- YES, he is exposed to little information
- NO, he never refers to external sources
- He mentions 5 "UNIVERSL EMOTIONS" (happy, sad, discuss, surprised, ...) **CONSERVATIVE FILTERS TOO!!!!**



PICKY

- 8, uses current filters a lot. **Like it because it reminds him filters in music stream services**
- Yes, very intuitive. **Suggests replace sliders with dropout values**
- YES, adores the cleanliness
- YES, he would definitely go for it
- **CONSERVATIVE FILTERS TOO!!!!**

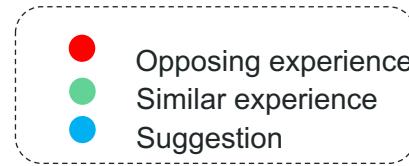
Learnings from Alpha Release: THE BIG QUESTION

Questions

- How much time do you think this product could save you?

IDEAL

- He agrees it can save him some time because it can contain additional and unique genres.



PICKY

- Definitely would save him time but express desire for the conventional filters.

NEXT ACTIONS (based on the feedback)

- Incorporate conventional genres
- Change sliders to dropdown
- Improve labeling of the emotional genres
- Need to find the distribution of the picky ones for appropriate business planning
- Find a way to measure success metric

What needs to be improved?

The screenshot shows a web browser window displaying a movie recommendation application. The URL in the address bar is <https://alphafilters-env.eba-pwmuxmfb.us-east-1.elasticbeanstalk.com/predict>. The page content includes:

- Header: "You are viewing Het Barot's screen" and "View Options".
- Score summary:
 - Realistic Settings Score = 4
 - Experience Excitement Score = 4
 - Fun Score = 3
 - IMDB Minimum Score = 8
- Movies recommended for hetbarot.nmims@gmail.com:
- Main heading: "Get the best movie recommendations for yourself!"
- Text: "Some explanation and instruction should be here. On the other hand, it's demanded to keep it tidy and clean. Can we come up with something very intuitive????"
- Form fields:
 - Username: "tipo" (highlighted with a red circle)
 - Slider for "Proud & Deep (between 1 and 5)": The value is set to 5 (highlighted with a red circle).
 - Slider for "Entertaining Music (between 1 and 5)": The value is set to 5.
 - Slider for "Realistic Settings (between 1 and 5)": The value is set to 5.

At the bottom of the browser window, there are standard controls: "Show all", "Leave", "Mute", "Start Video", "Participants", "Chat", "Share Screen", "Record", "Reactions", and "Leave".