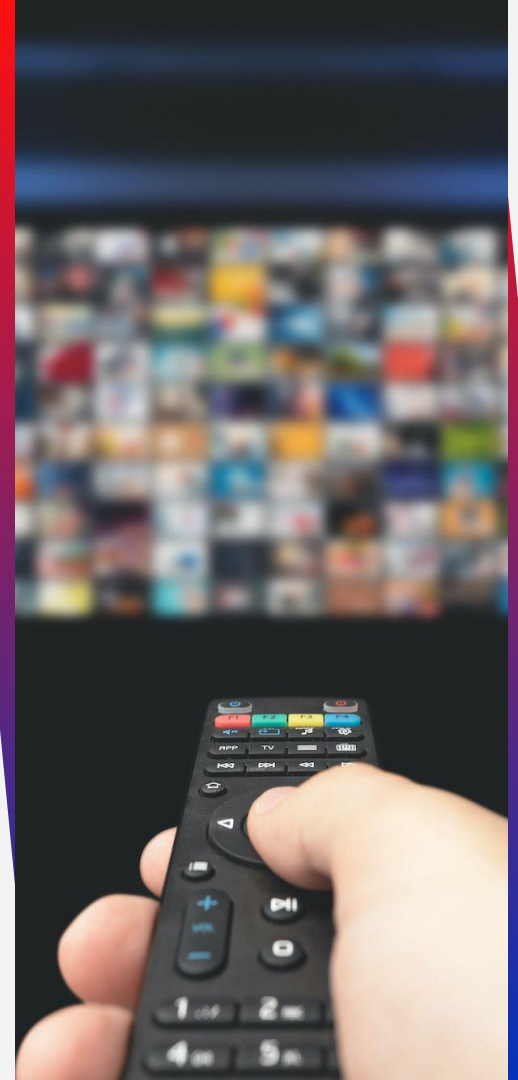# STREAMING PLATFORMS

NETFLIX VS **DISNEY+**

By Nareerat Moottatarn and Arkira Wongsim

TABLE OF CONTENTS

01.

PROBLEM STATEMENT

PROBLEM STATEMENT

Using Reddit posts we intend to perform sentiment analysis to understand and compare user attitudes towards Netflix and Disney+. Moreover, we aim to build a model to predict whether a streaming-related post is discussing Netflix or Disney+.

SCOPE OF STUDY

Data from posts in 2 Subreddits

r/netflix (535 posts)
Unofficial Netflix discussion, and all things Netflix related!
r/DisneyPlus (930 posts)
A subreddit for discussion of Disney's streaming service

NLP MODEL

To classify which post belongs to which Subreddit

SENTIMENT ANALYSIS

To compare positive or negative sentiments of
Netflix and Disney posts

02.

BACKGROUND

Streaming platforms are

on-demand online entertainment sources for TV shows, movies, music and other streaming media.

They are not only subscription-based streaming platforms like the Netflix or Spotify but also includes social media platforms that are free to publish content and free to watch.

Types of Streaming Platforms

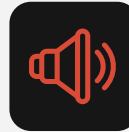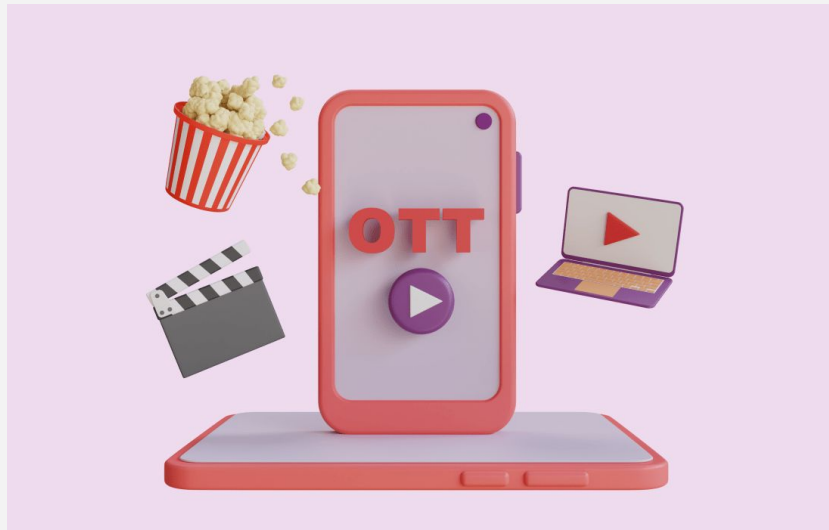| Social Media Platforms	OTT Platforms	Streaming Delivery Platforms
YouTube	Netflix	Roku
TikTok	hulu	tv
Facebook	Disney+	firetv

OTT PLATFORMS

Over-the-top platform refers to the technology that is delivers streamed content directly to user's internet-connected devices without the need for cable or satellite subscription.

Audio streaming
Spotify, Apple music

Messaging
Slack, Telegram, Whatsapp

Video streaming
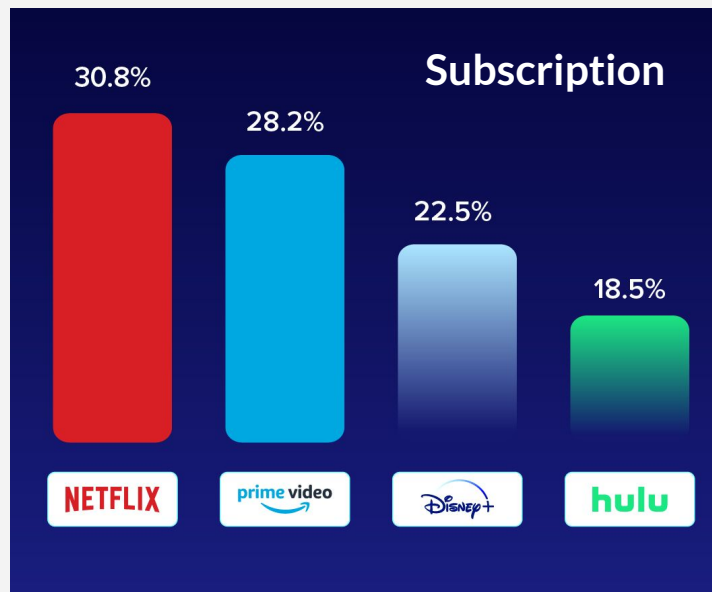Netflix, Amazon Prime, HBO Max, Disney+, Hulu

Video-streaming Industry

- Streaming services account for **40.3% of total television usage.**

- As of 2024, **1.4 billion people** uses video streaming services worldwide.

- In 2024, the video streaming market worldwide is estimated to reach **$108.50 billion.**

Subscription

30.8%

28.2%

22.5%

18.5%

NETFLIX

prime video

Disney+

hulu

NETFLIX

Disney+

	NETFLIX	Disney+
Launch Year	2007	2019
Subscribers globally	260.28 million	150.2 million
Geographic Availability	Over 190 countries	60 countries
Content	Mix of mainstream and Original content	Exclusive IP (Disney,Pixar, MARVEL,National Geographic) and Original content
Library Titles	Over 17,000 Shows and Movies	Around 15,000 Shows and Movies
Plans & Pricing (Thailand)	**Premium** 419 Baht/Month (4 Devices)	**Premium** 289 Baht/Month (4 Devices)

03.

ANALYSIS

Conversation Topics on Reddit
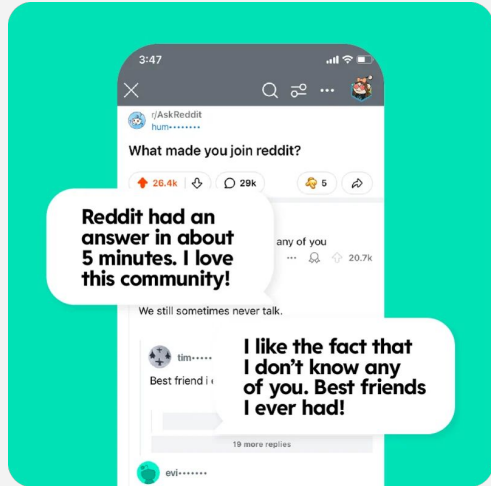
Number of Comments

The community comments on posts. Comments provide discussion.
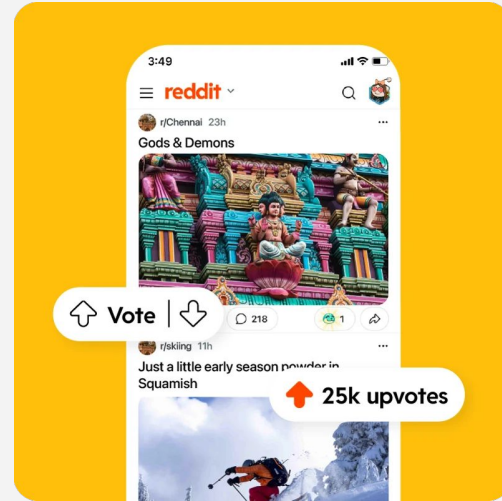
Score (Vote)

Comments & posts can be upvoted or downvoted. The most interesting content rises to the top.

Number of Comments
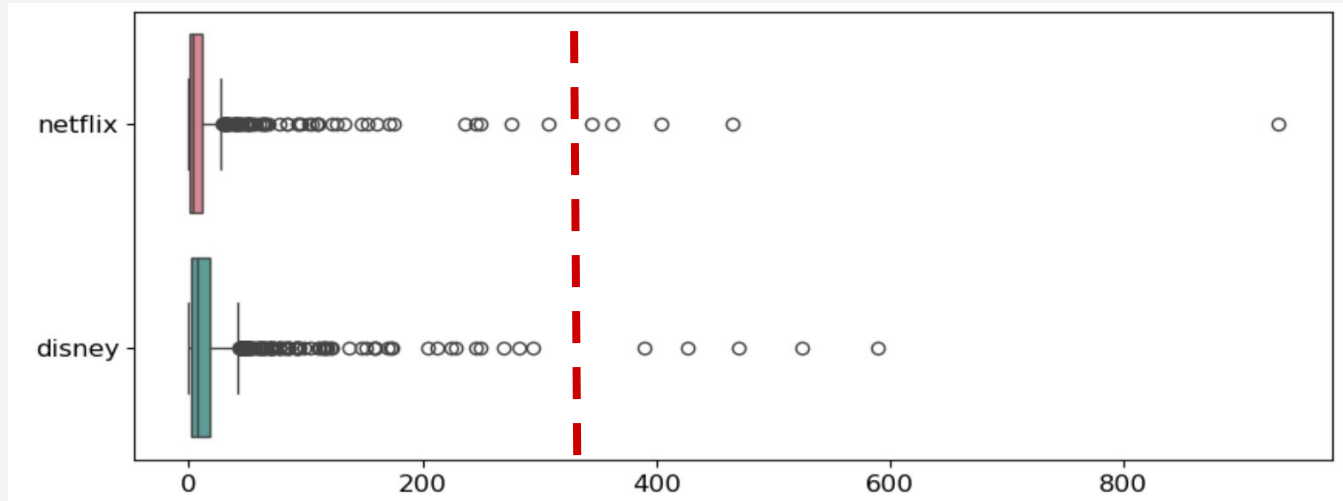
Looking at The Outliers (Top 5)

Number of Comments

Top 5 the most comments Netflix subreddit

Title	Comments
Outer banks new season has crippling audio issues	932
Menendez brothers were molested and should walk free, says prosecutor	465
Sweet Bobby: My Catfish Nightmare	404
Just finished Alien Abduction in Manhattan. I'm on Carol's side.	362
time cut didn't make any sense	345

Issues

After-The-Movie Conversations

Number of Comments

Top 5 the most comments Disneyplus subreddit

Title	Comments
Describe your favorite Disney movie only using emojis🤴🦎☀️!',	590
???	525
What's a movie everyone keeps insisting is great but you just don't get the hype?	472
That's just obscene	427
Disney+ terms prevent allergy death lawsuit, Disney says	390

Increased Price

Movie preferences

Score (Vote)
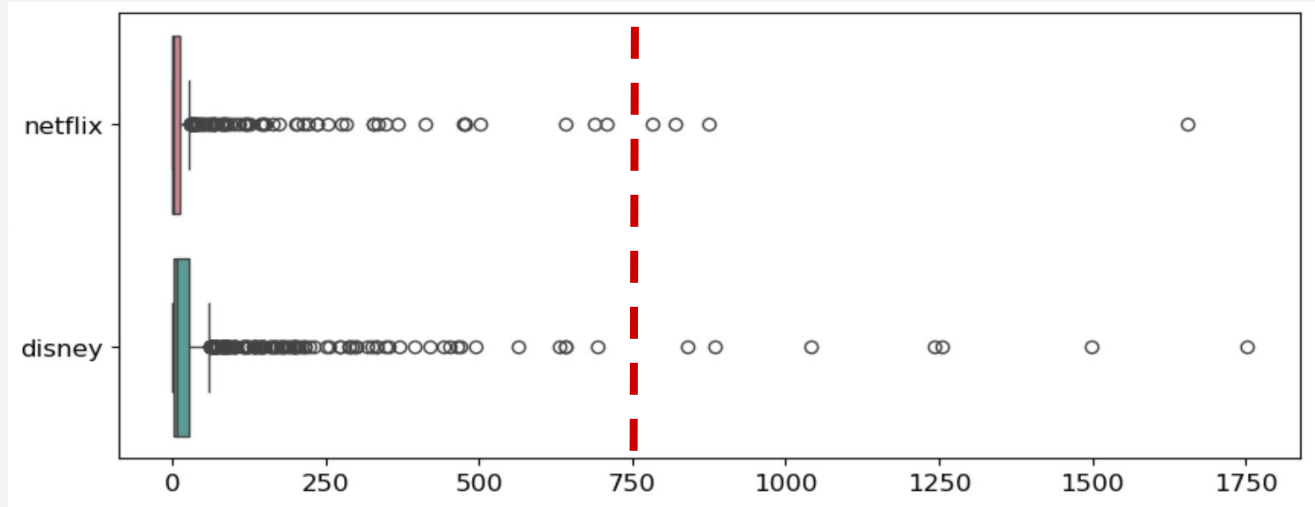
Looking at The outliers (more than 750 score)

Score (Vote)

Post on Netflix subreddit with more than 750 score

Title	Score
I just rewatched Mindhunter. What a colossal blunder to let Fincher walk.	1,656
Menendez brothers were molested and should walk free, says prosecutor	875
The need for growth is ruining Netflix	821
The Queen's Gambit debuted 4 yrs ago today on Netflix	783

After-The-Movie Conversations

Operational Issue

Score (Vote) :

Post on Disneyplus subreddit with more than 750 score

Title	Score
???	1,751
That's just obscene	1,499
Agatha All Along' Falls Out of Top 10 Streaming Shows Last Week Despite Significant Plot Twist"	1,254
Children's PG profile	1,242
Not sure I can afford it anymore…	1,042
'Phineas and Ferb Revival - First Poster',	884
Disney+ has just removed all the Indiana Jones movies directed by Steven Spielberg from the platform worldwide	840

Increased Price

After-The-Movie Conversations

2 Models
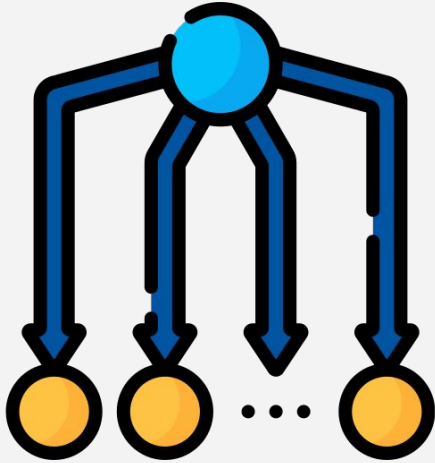
MODEL 1 : NAIVE BAYES
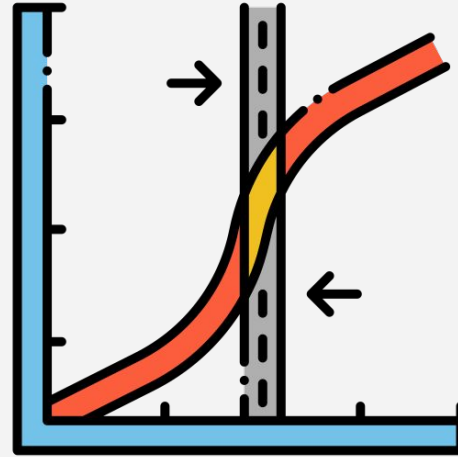
MODEL 2: LOGISTIC REGRESSION

Baseline

Features: We used the 'title' of the posts to train and test the data. We chose to stratify y and split with size 80/20, and random state for reproducibility.

Disney (0) : 63.48 %	
Netflix (1) : 36.52 %	

Preprocessing

1. Removed special characters
2. Removed 'Disney' and 'Netflix'
3. Spelling mistakes such as "fo" to "for"
4. Lemmatize
5. Tokenize
6. Stopwords: "all_stops" because lemmatizing created new stop words in addition to the original stop_words='english'

MODEL 1 : NAIVE BAYES

Reason : MultinomialNB is used since the columns of X are all integer counts.

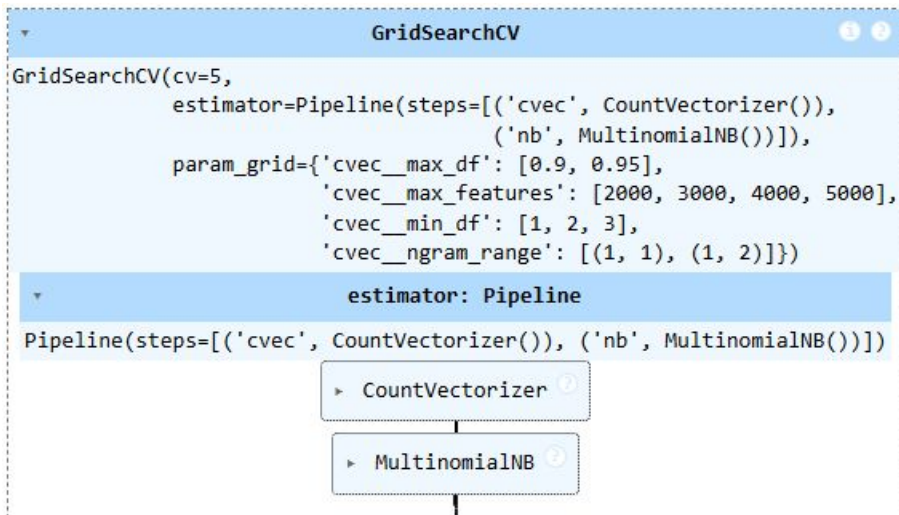**cvec = CountVectorizer(max_features=500, tokenizer=lemma_tokenize)**

Model

```
                              GridSearchCV                        ①  ②
GridSearchCV(cv=5,
             estimator=Pipeline(steps=[('cvec', CountVectorizer()),
                                        ('nb', MultinomialNB())]),
             param_grid={'cvec__max_df': [0.9, 0.95],
                         'cvec__max_features': [2000, 3000, 4000, 5000],
                         'cvec__min_df': [1, 2, 3],
                         'cvec__ngram_range': [(1, 1), (1, 2)]})
```

```
                          estimator: Pipeline
Pipeline(steps=[('cvec', CountVectorizer()), ('nb', MultinomialNB())])

                          ▸ CountVectorizer ⓘ

                          ▸ MultinomialNB ⓘ
```
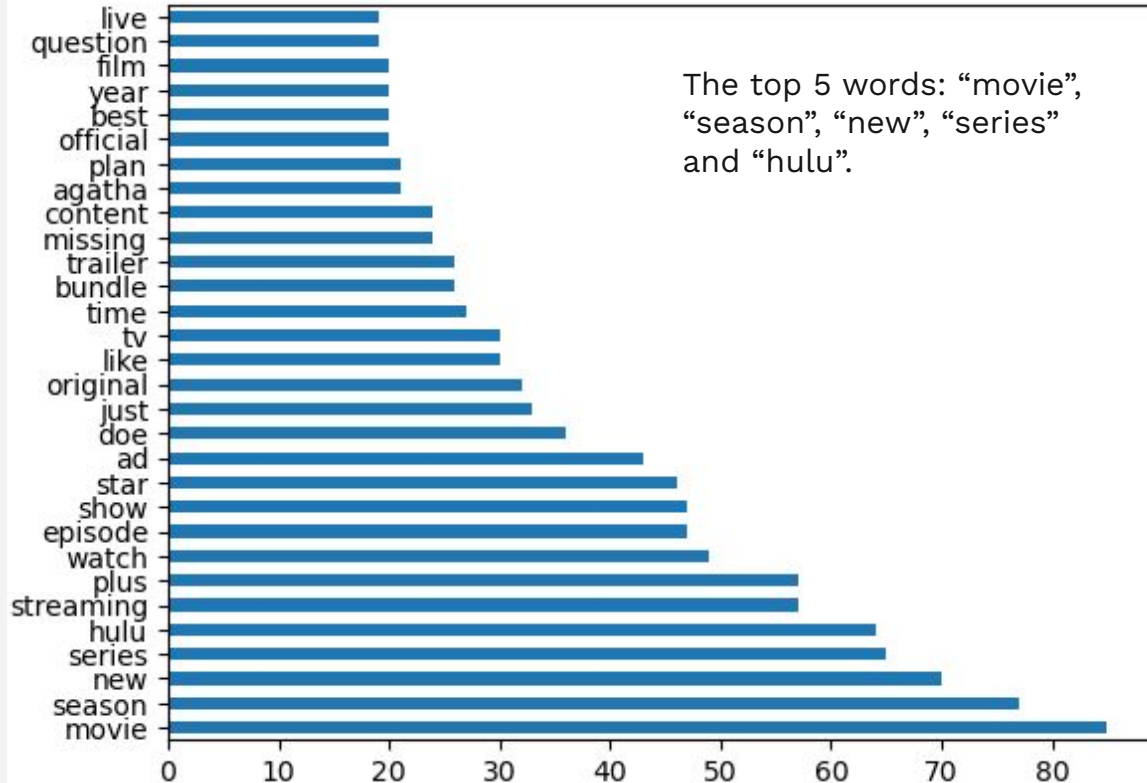
Best Params

```
gs.best_params_

{'cvec__max_df': 0.9,
 'cvec__max_features': 5000,
 'cvec__min_df': 1,
 'cvec__ngram_range': (1, 2)}
```

Accuracy

Train Score : 72%

Test Score : 78%

Top 15 Common Words

The top 5 words: "movie", "season", "new", "series" and "hulu".

MODEL 2: LOGISTIC REGRESSION

Reason : Logistic regression is the most common binary classification algorithm and interpretable.

Default
- **Default penalty='L2'**
- **Using X training data that was CountVectorized**

Accuracy
Train Score : 85%
Test Score : 76%

GRIDSEARCH
- **The accuracy was score decreased.**
- **Best Score: 0.70**
- **Best Params {'C': 1.0, 'penalty': 'l2', 'solver': 'liblinear'}**

Results

Model	Algorithm	Evaluation	Data	Result
1	NB			0.72
2	LR (Default : L2 regularization)	accuracy	netflix, disney	0.85
3	LR (Gridsearch)			0.70

Model Improvements

1. More data
> sample size increase

2. Explore other Logistic Regression
> hyperparameters (penalty, C, solver, max_iter, tol, l1_ratio)

3. Explore other CountVectorizer
> hyperparameters (max_df, min_df, max_features, ngram_range)

4. Explore other classification models such as kNN, Random Forest.

SENTIMENT ANALYSIS

r/netflix · 12 days ago
Crashingpigon15

Topic

Outer banks new season has crippling audio issues

Body

The end of season 4 has just come out, and I was watching it. Episodes 8,9, and 10 have major audio issues. After the Netflix "da dum" at the beginning of the episode, there is a second one, and that is followed by there being no audio on the actual show. The problem is only with these specific episodes of outer banks, the other episodes are fine, and other shows/movies are fine. I have restarted my TV, restarted Netflix all to no avail. Anyone have an idea of when this will be fixed?

| Negative <= −0.05 | Neutral (−0.05,0.05) | Positive >= 0.05 |

SENTIMENT ANALYSIS

| Negative <= -0.05 | Neutral (-0.05,0.05) | Positive >= 0.05 |

Title	Avg. Score	Negative	Neutral	Positive
Disney+	0.0428	17%	58%	25%
Netflix	-0.0164	25%	52%	23%

The sentiment of 'titles' is slightly negative for Netflix compared to Disney.

Body	Avg. Score	Negative	Neutral	Positive
Disney+	0.1336	21%	41%	38%
Netflix	0.1073	28%	30%	42%

Both platforms have positive sentiments in the 'body'.

04.

CONCLUSION

Conclusion

Model performed best

Naive Bayes / Logistic Regression

Content

The top 5 words: "movie", "season", "new", "series" and "hulu".

Sentiment

Both Netflix and Disney have positive 'body' post sentiments, but Netflix with slightly negative 'title' sentiments.

Thank you