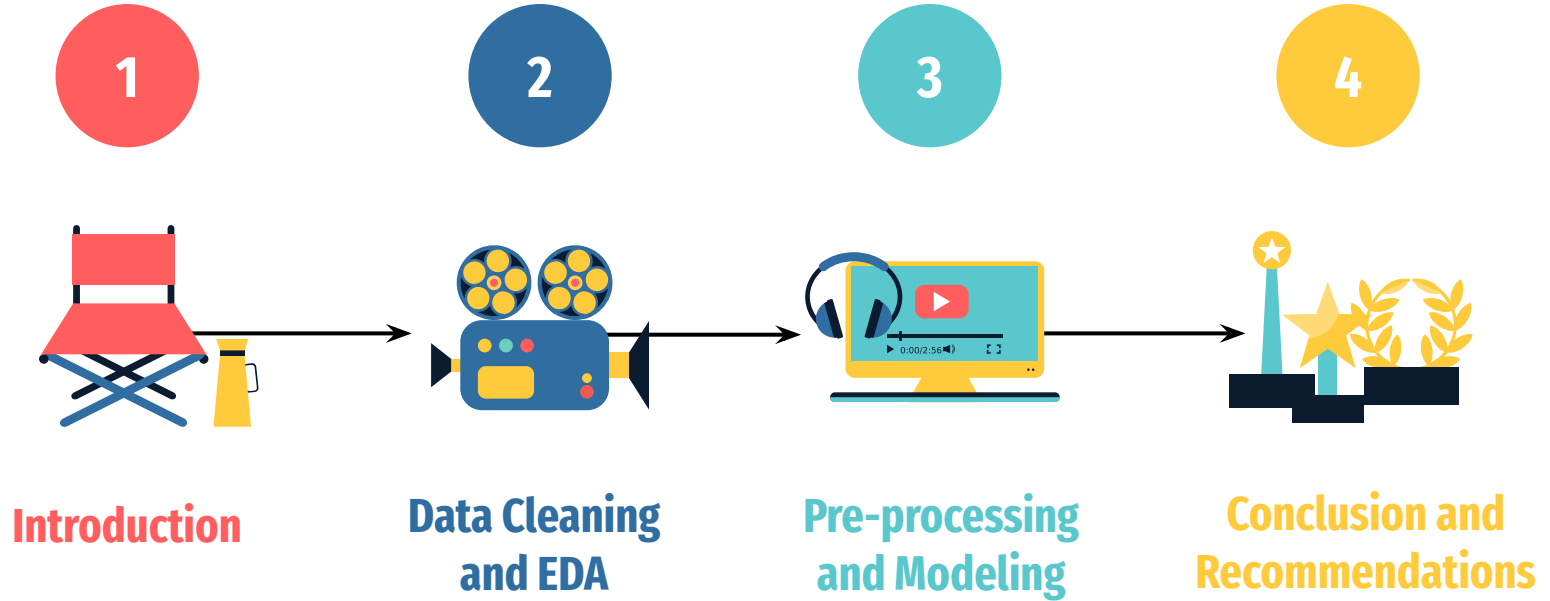


Netflix vs. Disney+

Classification Project



CONTENT



01

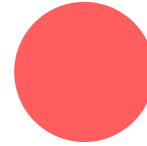
Introduction

Background & Problem Statement

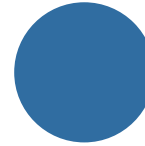


BACKGROUND

What happen to Netflix?



1 million subscribers loss



General fears about the company



Competition with company like Disney +, HBO

PROBLEM STATEMENT



Build a **classification model** contrasting Netflix and Disney Plus, then evaluate the **significance of the features** to offer suggestions on how to **boost Netflix's subscription**.

02

Data Cleaning and EDA



DATA COLLECTION

**Use Pushshift
API:**

To scrape reddit
posts

2 subreddits:

r/netflix
r/DisneyPlus

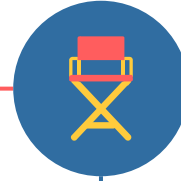
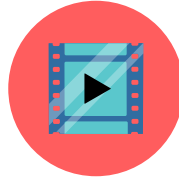
**1300 submission
posts**

for each subreddit



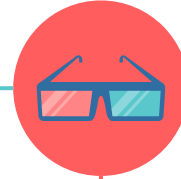
DATA CLEANING

**Combine selftext +
title**



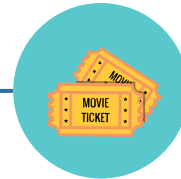
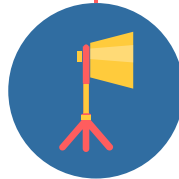
**Drop missing values/
duplicates**

**Deal with deleted
and removed posts**



**Remove special
characters**

Lemmatize tokens

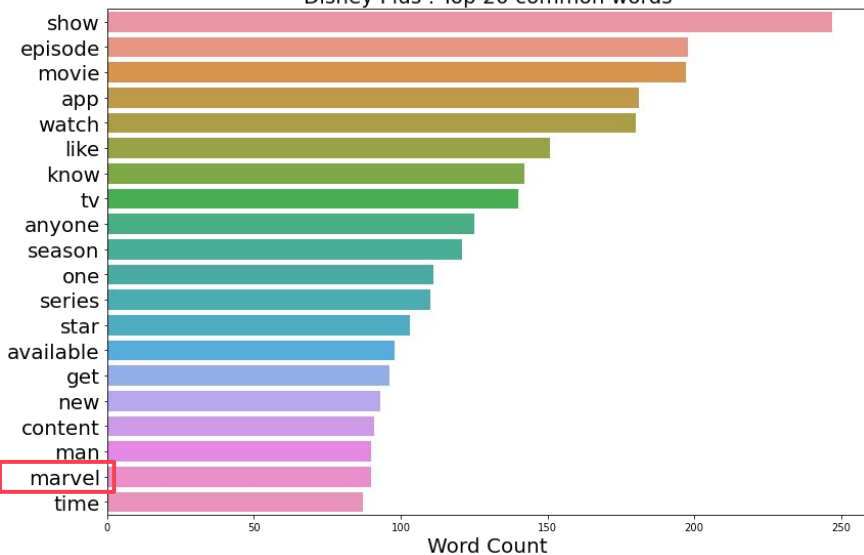


**Add additional stop
words : Netflix,
Disney Plus**

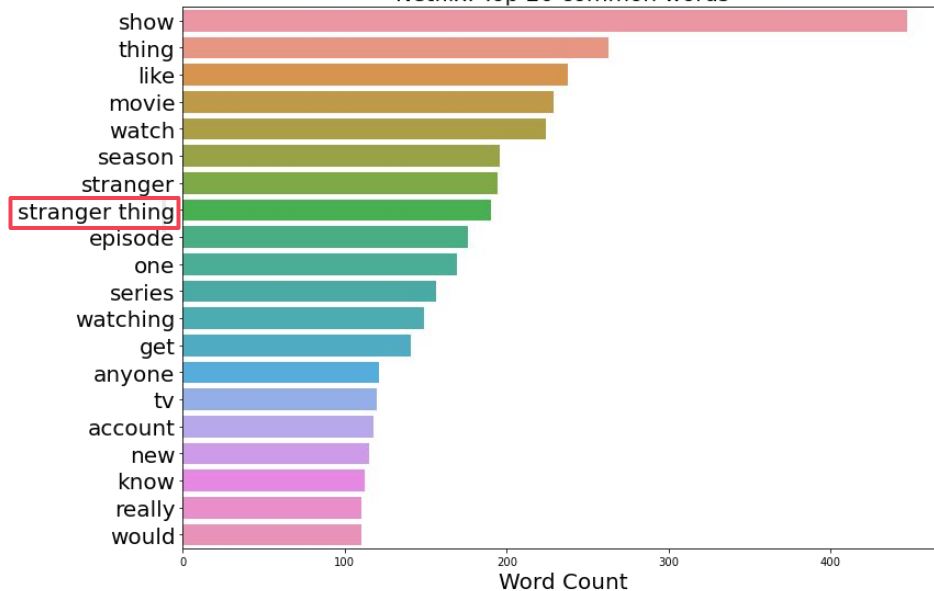


TOP COMMON WORDS

Disney Plus : Top 20 common words



Netflix: Top 20 common words



03

Modeling

Preprocessing & Evaluation



MODEL SUMMARY

Best model based on:



Highest Accuracy



Least Overfitting

	model	vectorizer	best_score	train_score	test_score(Accuracy)
0	Logistic Regression	Count Vectorizer	0.746867	0.987994	0.769115
1	Logistic Regression	TF-IDF Vectorizer	0.763367	0.957479	0.760120
2	Multinomial NB	Count Vectorizer	0.785875	0.964482	0.775112
3	Multinomial NB	TFI-DF Vectorizer	0.775381	0.942471	0.770615
4	Random Forest	Count Vectorizer	0.721845	0.864432	0.713643
5	Random Forest	TFI-DF Vectorizer	0.731846	0.889445	0.734633

An illustration of three people sitting on the ground, looking at the model summary table. A woman with dark hair is on the left, a man with short dark hair is in the middle, and a woman with long dark hair is on the right. They are all wearing casual clothing. The background shows stylized green bushes and a blue sky.

TOP FEATURES

NETFLIX

- Stranger Thing
- Anime
- Lucifer
- Resident Evil
- Umbrella Academy



DISNEY+

- Marvel
- Lightyear
- Spider Man
- Obi Wan Kenobi
- Star Wars

04

Conclusion

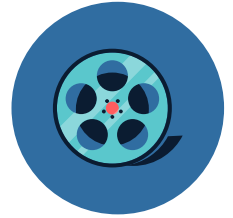
Recommendation and Future
Improvements



Business Recommendations for Netflix



**Maintain its wide range
movie selections**



**Improve the quality of
movie productions while
also maintaining the
subscription price**



FUTURE IMPROVEMENTS



1

Collect more data points

2

Include more relevant or similar subreddits

3

Explore additional features (i.e. comments)

4

Add more specific stop words

THANKS!

Any questions?

