Movie Recommender System

Priscilla
26-01-2023

Content


ROLL: 01
TAKE: Introduction

Recommender Systems

Movie Streaming Service

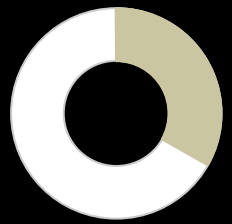
Netflix

E-Commerce

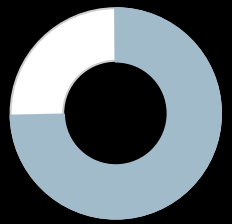
Amazon

amazon

Benefits to Businesses/Consumer



35% Amazon



75% Netflix

McKinsey

35% of what consumers purchase on Amazon and **75%** of what they watch on Netflix comes from product recommender systems

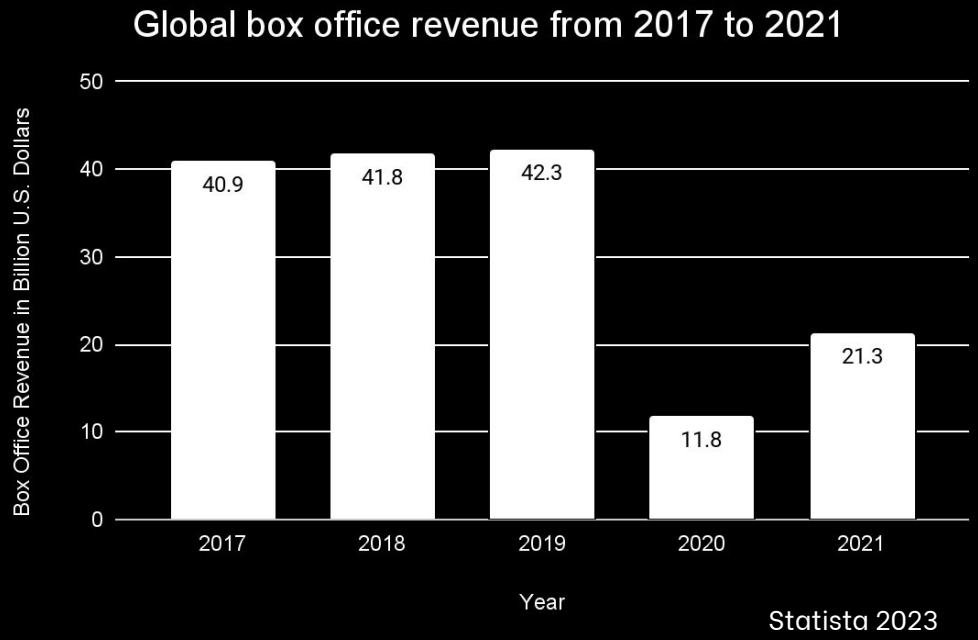
Benefits to Consumers/Businesses

Customer Satisfaction

Personalization/Discovery

Revenue

Film Industry



- Steadily **increasing** over the years before COVID-19 in 2020
- However, revenue is in an **upward trend** since 2020

Consumers

- Fast
- Less Effort

Problem Statement

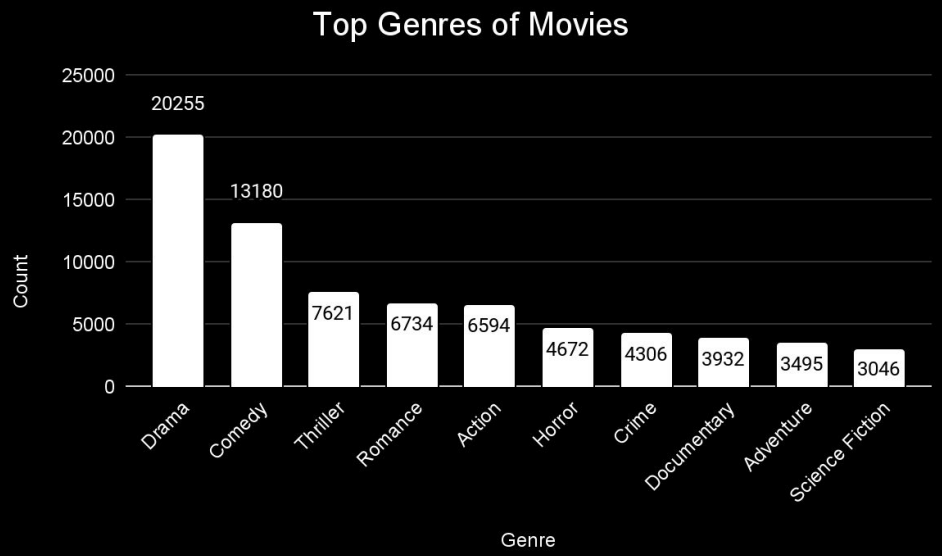
The aim of this project is to build various types of movie recommender systems and evaluate these systems based on their performance.


ROLL:
02
TAKE:
EDA

Dataset

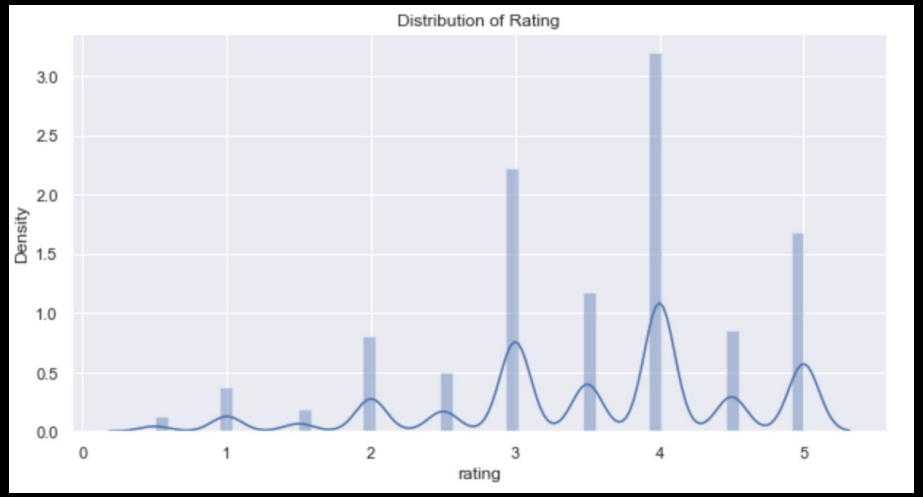
Metadata	Metadata of all 45,000 movies released on or before July 2017
Keywords	
Credits	
Ratings	26 million ratings from 270,000 users for all 45,000 movies

Genre



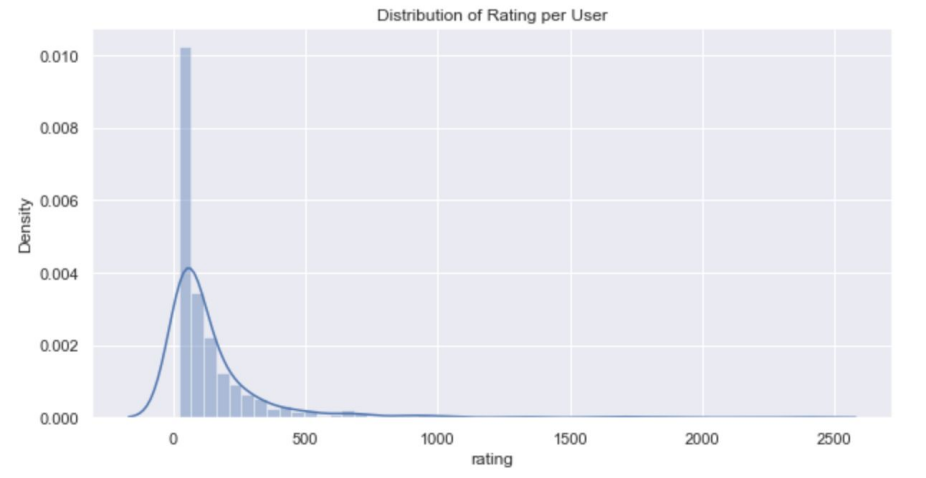
- Top genres of movies in the dataset are **Drama**, **Comedy** and **Thriller**

Ratings



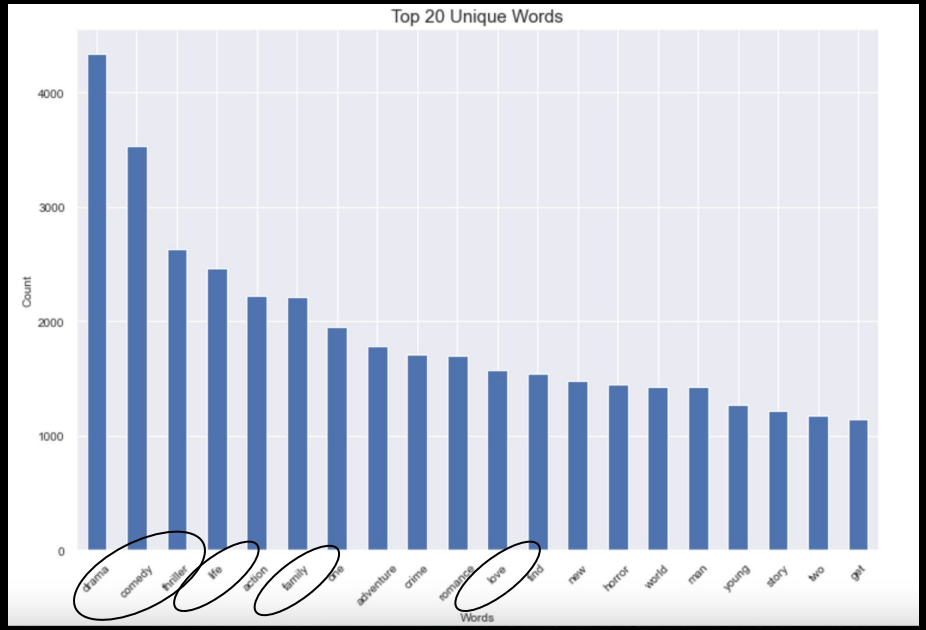
- Majority of user gave the movies a **3-4** out of 5 rating

Ratings per User

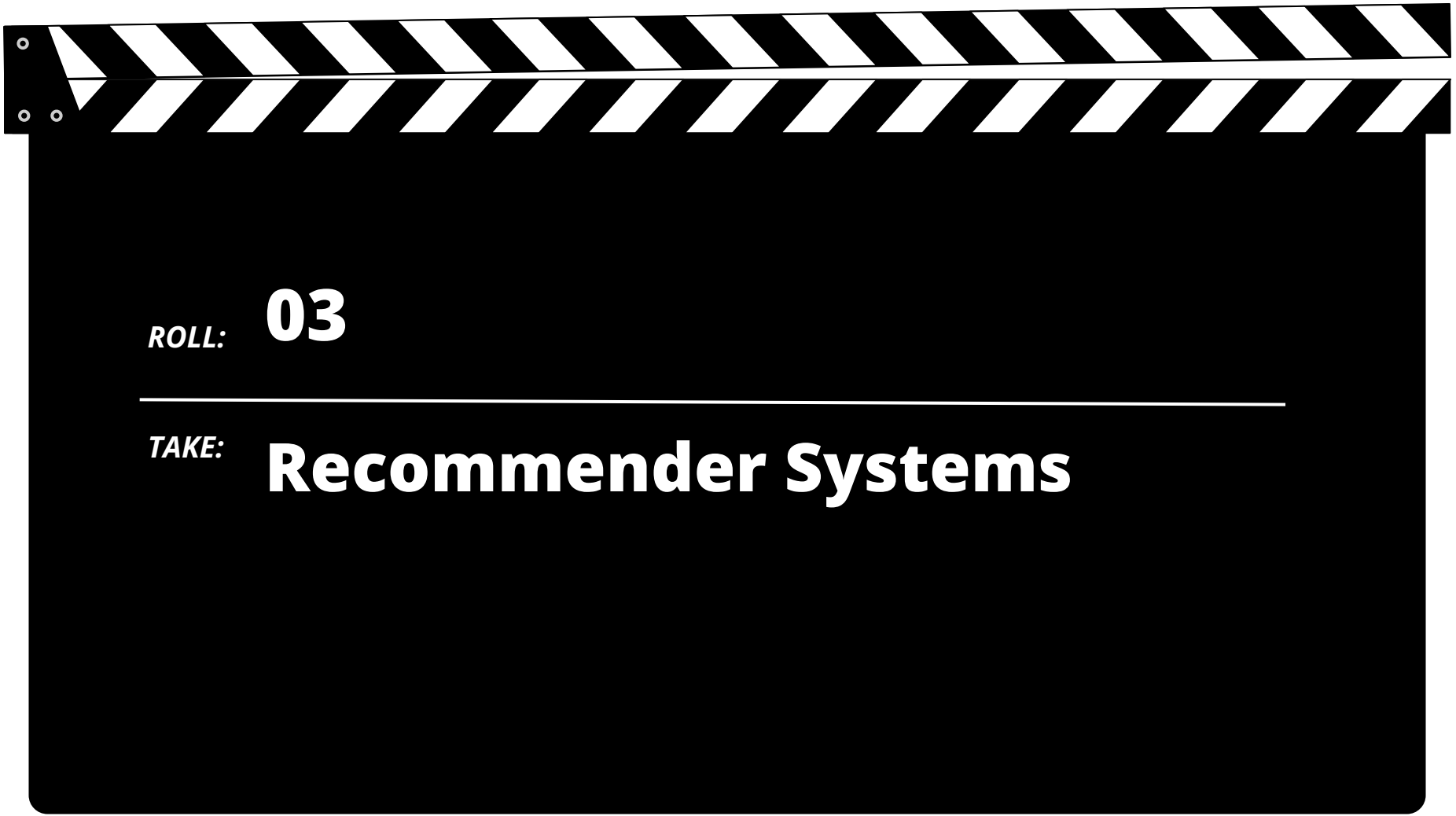


- Each user rates a minimum of **20** movies and up to a maximum of **2391** movies.

Top 20 Unique Words



- Using the words from movie title, overview, tagline, genres, keywords and cast
- Top words are **Drama**, **Comedy** and **Thriller** which coincides with the top 3 genres of movies
- Movies are based on the themes of **Life**, **Family** and **Love**.


ROLL: 03
TAKE: Recommender Systems

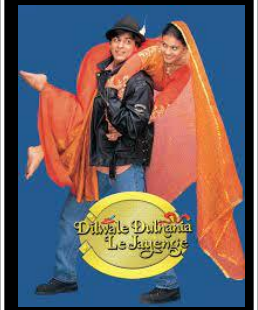
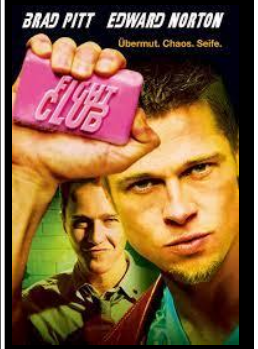
Basic Recommender System

IMDB Weighted Rating Formula

Weighted Rating (WR) = v/(v+m.R) + m/(v+m.C)

where,
v is the number of votes for the movie
m is the minimum number of votes required to be listed in the charts
R is the average rating of the movie
C is the mean vote across the whole report

Top 5 movies - WR

	1	2	3	4	5
Movie					
WR	8.3	8.0	8.0	8.0	8.0

Top 5 movies - WR and Genre

	1	2	3	4	5
Movie					
WR	7.9	7.9	7.9	7.8	7.8

Content-Based Filtering

***STEP:* 01**

Select the Categories of interest

Title, Overview, Tagline, Genres, Keywords, Cast

***STEP:* 02**

Merge all columns into one column

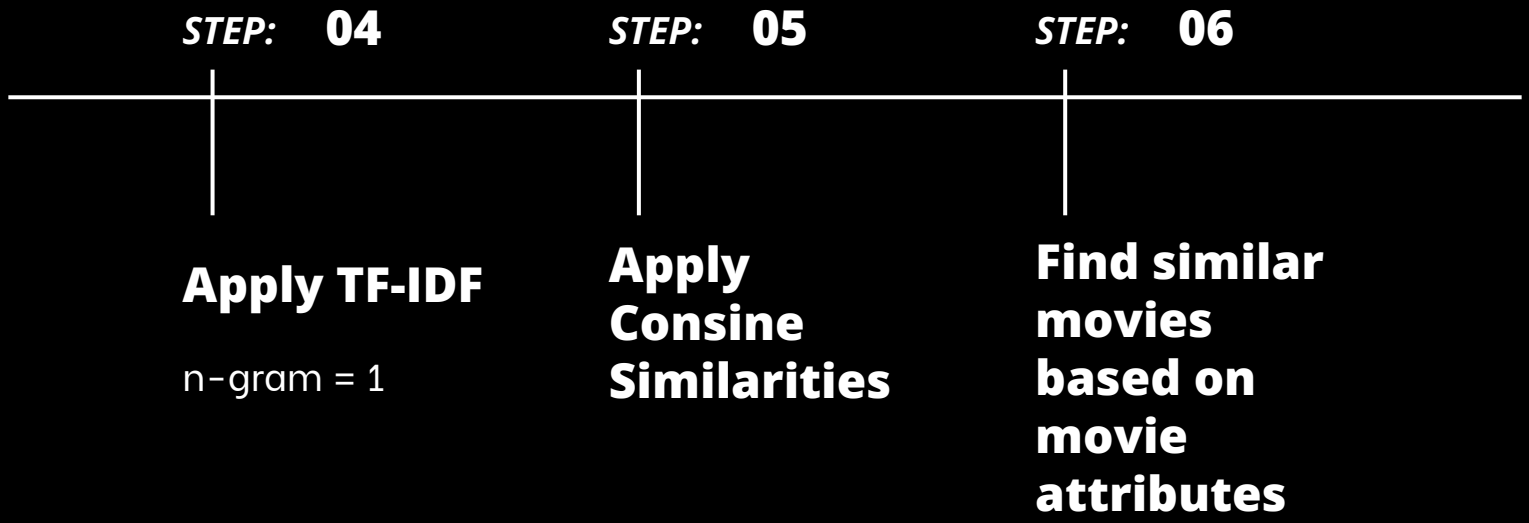
tags

***STEP:* 03**

Clean Text

Remove punctuations, stopwords and lemmantization

Content-Based Filtering



Content-Based Filtering

Minions

1	Mower Minions (0.413)
2	Minions: Orientation Day (0.362)
3	Banana (0.273)
4	Despicable Me 2 (0.251)
5	Despicable Me (0.215)

The Dark Knight

1	The Dark Knight Rises (0.307)
2	Batman Begins (0.273)
3	Batman Returns (0.270)
4	Batman Forever (0.257)
5	Batman: Under the Red Hood (0.229)

Collaborative Filtering

Collaborative Filtering

Filters movies a user might like based on the ratings given by similar users

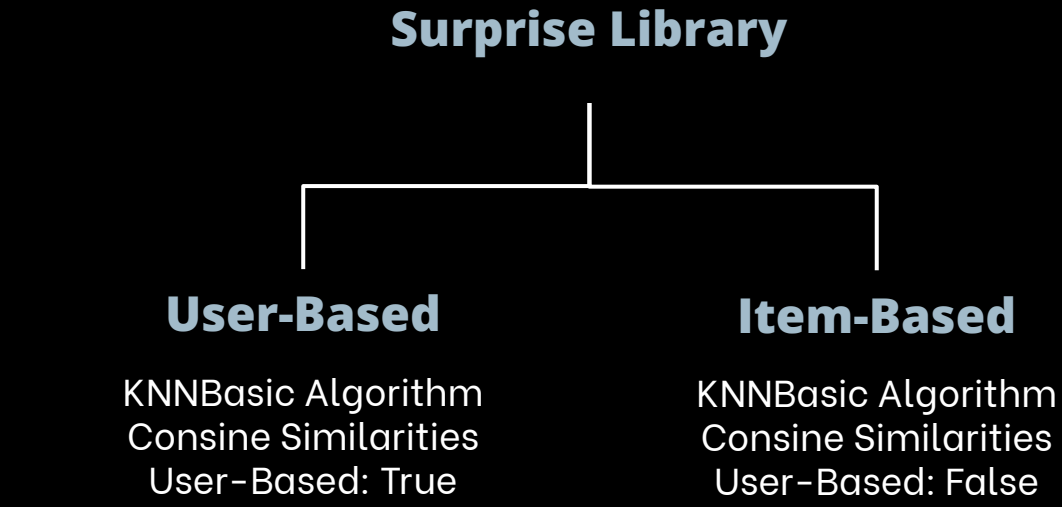
Memory-Based

User-Based
Item-Based

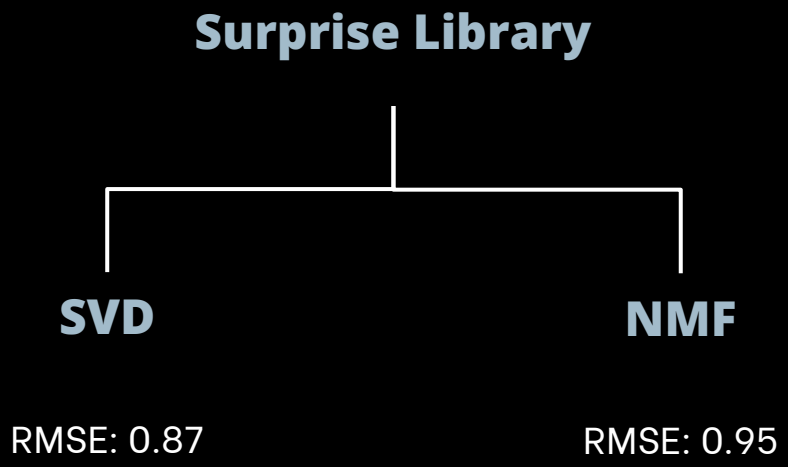
Model-Based

Matrix Factorization based on SVD and NMF

Memory-Based Collaborative Filtering



Model-Based Collaborative Filtering



User 671

Mystery/Thriller

Drama/Comedy

Action/Sci-Fi

The Poseidon Adventure
The Million Dollar Hotel
Terminator 3
The Searchers
Boogie Nights
Jacob’s Ladder

User 671

User-Based
Blown Away
Rio Bravo
The Celebration
Stalag
Gentlemen Prefer Blondes

Item-Based
2046
The Protector
Everything is Illuminated
K-PAX
The Silence of the Lambs

User 671

SVD
Nell
Flags of Our Fathers
Fools Rush In
Shriek If You Know What I Did Last Friday the Thirteenth
Sleepless in Seattle

NMF
Gentlemen Prefer Blondes
Once Upon a Time in Mexico
Ninotchka
Speed Racer
Backdraft

ROLL: **04**

TAKE: # Conclusion

Conclusion

	Advantage	Disadvantage
Basic	• Simple • Easy to build	• Not unique to user preference • Lack of personalization
Content-Based Filtering	• Recommendations are specific to user, do not require other user data, can scale to large number of users	• Cold Start • Recommendations are based on existing interest of user

Conclusion

	Advantage	Disadvantage
Collaborative Filtering	• All users are taken into consideration and people with similar tastes and preference are used to suggest new movies to the primary user	• Sparsity • Scalability

Recommendations

1. Explore Deep Neural Networks for Collaborative Filtering
2. Explore Hybrid Content-Based and Collaborative Filtering Recommender Systems

THANKS!

ROLL:

DO YOU HAVE ANY QUESTIONS?

youremail@freepik.com
+91 620 421 838
yourcompany.com

TAKE:

SCENE: