

DOSEN PEMBIMBING
 HENDRA BUNYAMIN, S.SI., M.T.
 SEMESTER GANJIL 2021/2022

1873006 | SIENIE CELICIA DEWI
 1873006@MARANATHA.AC.ID

Netflix merupakan salah satu layanan streaming acara tv dan film yang populer dan telah memiliki lebih dari 204 juta pengguna. Dalam penelitian ini dilakukan eksplorasi terhadap dataset acara tv dan film Netflix melalui tahapan EDA (Exploratory Data Analysis) untuk mengetahui karakteristik dari data dan mendapatkan informasi dari data yang ada. Kemudian, dilakukan eksplorasi untuk pembuatan sistem rekomendasi dari suatu tayangan, clustering dan eksplorasi N-Gram

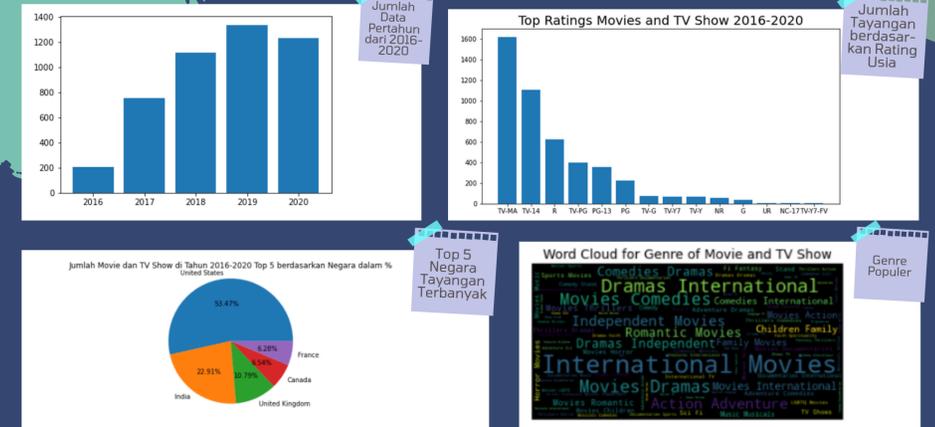
DATASET

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	TV Show	3%	NaN	João Miguel Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	2020	TV-MA Seasons 4	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	7:19	Jorge Michel Grau	Demian Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	2016	TV-MA 93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	2011	R 78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly...	United States	November 16, 2017	2009	PG-13 80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	2008	PG-13 123 min	Dramas	A brilliant group of students become card-coun...

```
df.dtypes
show_id      object
type         object
title        object
director     object
cast         object
country      object
date_added   object
release_year int64
rating       object
duration     object
listed_in    object
description  object
dtype: object
```

```
df.shape
(7787, 12)
```

HASIL EDA



HASIL SISTEM REKOMENDASI

HASIL CLUSTERING

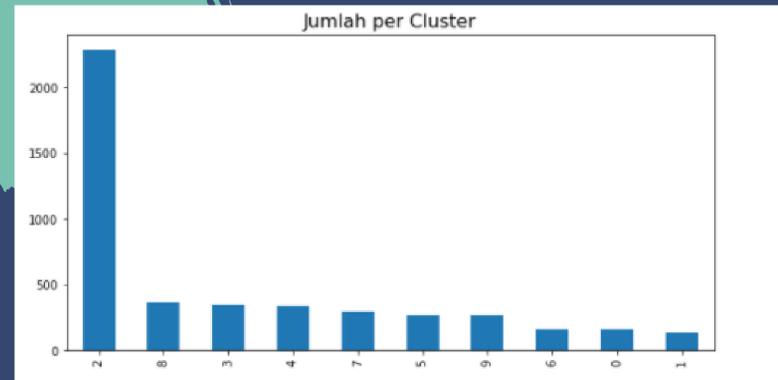
HASIL N-GRAM

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	features	ids
2866	s4753	Movie	Pandora	Jung Hyeon, Park Ji-hye, Kim So-hyun, ...	South Korea	2017	2016	TV-MA	127	Dramas, International Movies	When an earthquake strikes a coastal village near...	Pandora	2866

Count Vectorizer

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	features	ids
3021	s571	Movie	The Forgiven	Clare Danes, Jason Momoa, James Van Der Beek, ...	United States	2017	2014	TV-MA	89	Horror Movies	After a war-torn village is destroyed, a lone soldier...	The Forgiven	3021

TF-IDF Vectorizer



```
bigrams2 = nltk.bigrams(textng2)
freqb2 = nltk.FreqDist(bigrams2)
freqb2.most_common(5)

[[('base', 'true'), 16],
 [('fall', 'love'), 16],
 [('join', 'force'), 13],
 [('take', 'stage'), 13],
 [('world', 'war'), 12]]

trigrams2 = nltk.trigrams(textng2)
freqt2 = nltk.FreqDist(trigrams2)
freqt2.most_common(5)

[[('world', 'war', 'ii'), 6],
 [('must', 'decide', 'whether'), 4],
 [('travel', 'back', 'time'), 3],
 [('take', 'stage', 'brooklyn'), 3],
 [('game', 'cat', 'mouse'), 3]]
```

