HOW POPULAR WILL THIS VIDEO BE ON YOUTUBE ?

Team : Noura Al-Otaibi, Miatha Al-Qahtani, Alaa Al-Ghamdi


YouTube
TRENDING

Introduction

As Youtubers earn money through the advertising and bonus of their videos, increasing the popularity of their videos is a priority. This project aims to predict the behavior of the video that is going to be uploaded to YouTube. An equation is developed to manually classify all the videos into binary groups: non-popular and popular.

01 **NUMPY**

02 **MATPLOTLIB AND SEABORN FOR PLOTTING**

03 **Pandas for data manipulation**

04 **SCIPY**

05 **SKLEARN**

06 **WORDCLOUD**

Exploratory Data Analysis

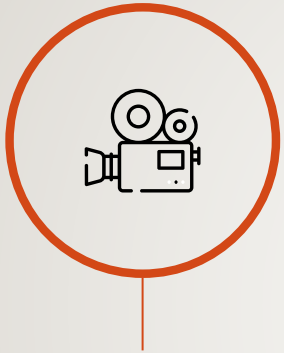
Statistical analysis

Data cleaning

Preparing data to modeling

Modeling and compare models

SDAIA ACADEMY



DATA SOURCE

Kaggle as an open-source.
we selected the trending videos of YouTube from the United States and Canada, consisting of 29089 unique videos.

DATASET

RECRDS: 81830
Columns: 17

SDAIA ACADEMY

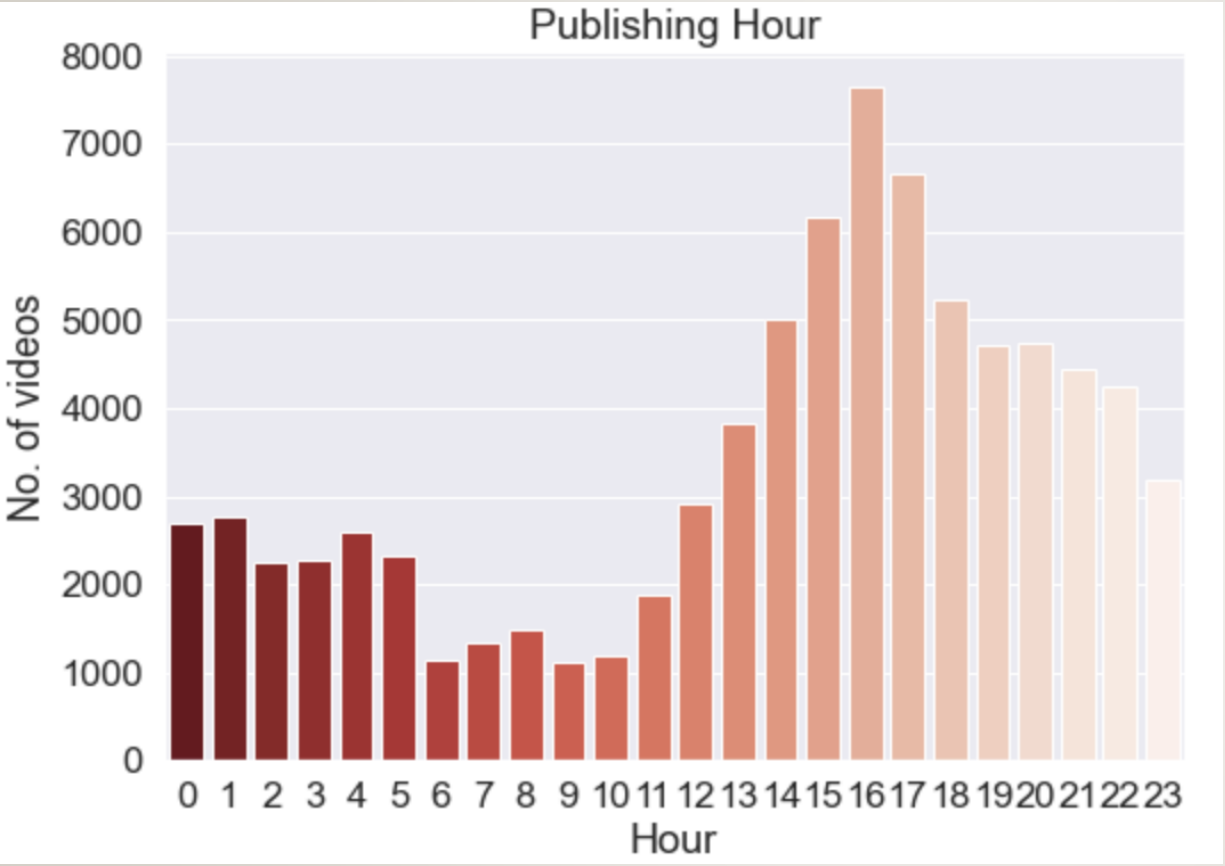
- **We created a new y for prediction by eqution**

$$Score = \frac{Comments}{Views} \times (Likes - 1.5 \times dislikes)$$

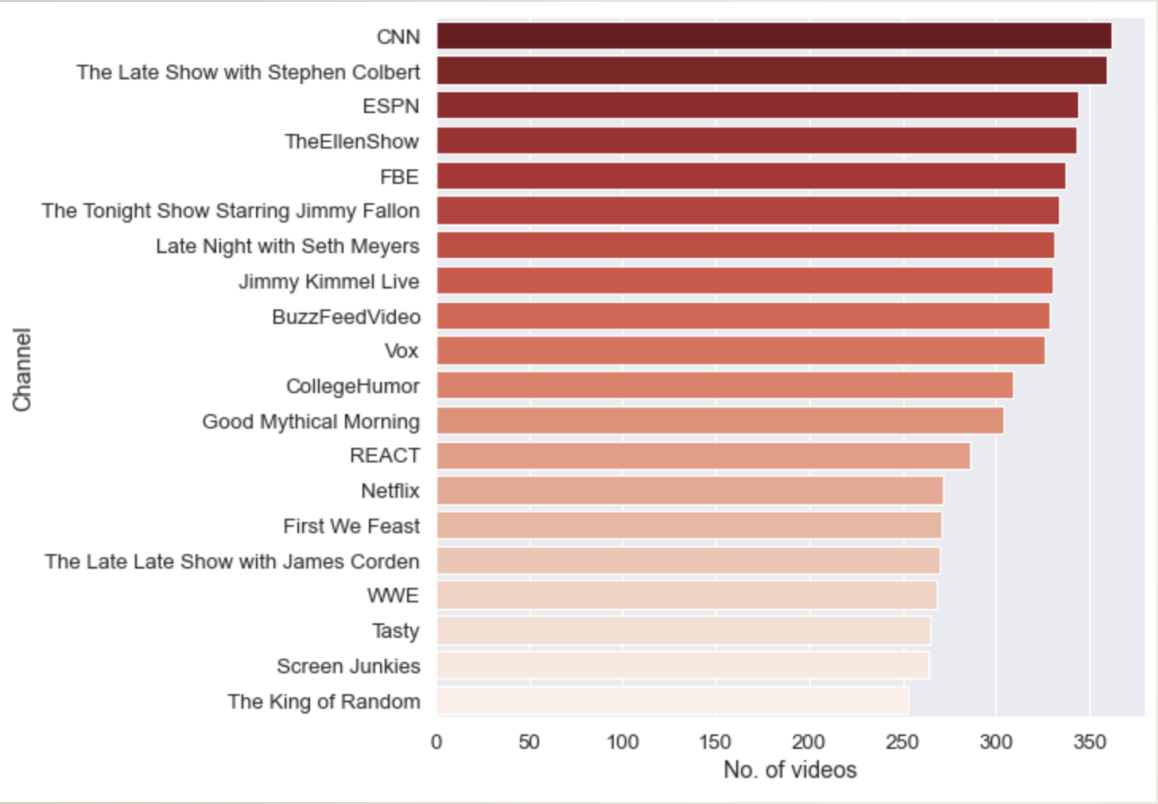
$$y = \begin{cases} 0, & Views < 100,000 \\ 1, & Views \geq 100,000 \quad Score < 0 \\ 2, & Views \geq 100,000 \quad 0 \leq Score < 300 \\ 3, & Views \geq 100,000 \quad Score \geq 300 \end{cases}$$

FEATURES ENGENDERING :

1. Score
2. Difference day between publish and trends
3. Tegs count
4. Video Type

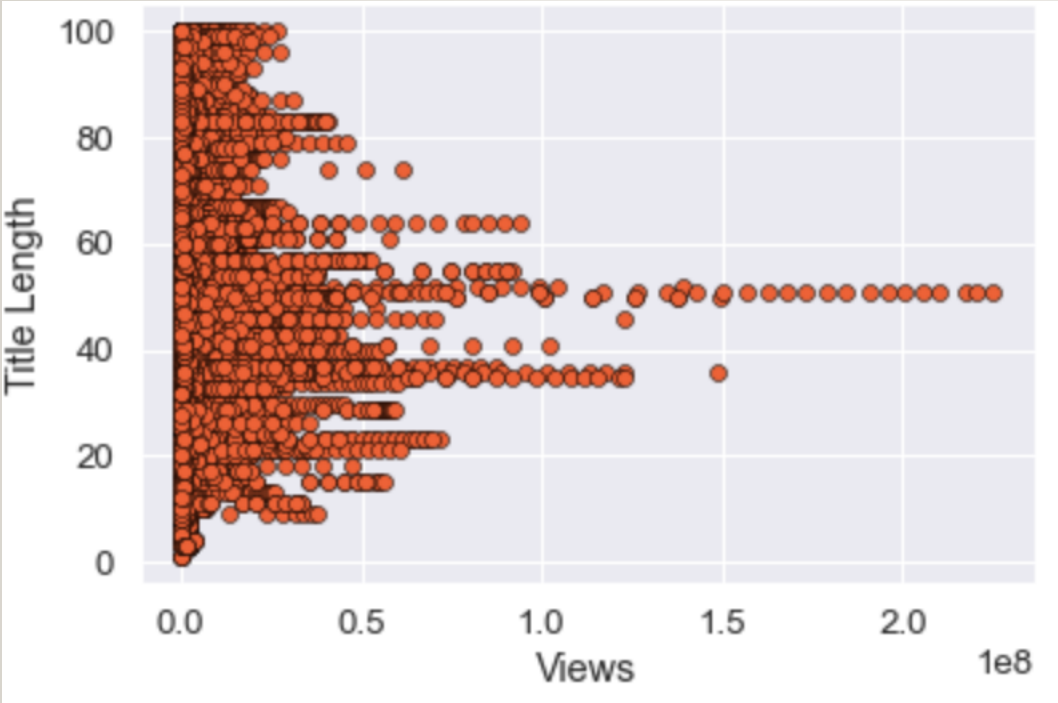


observation that number of videos published at 4 pm, 3 pm, 5 pm and 6 pm trend more than the number of videos uploaded 8 am, 7 am, 6 am.



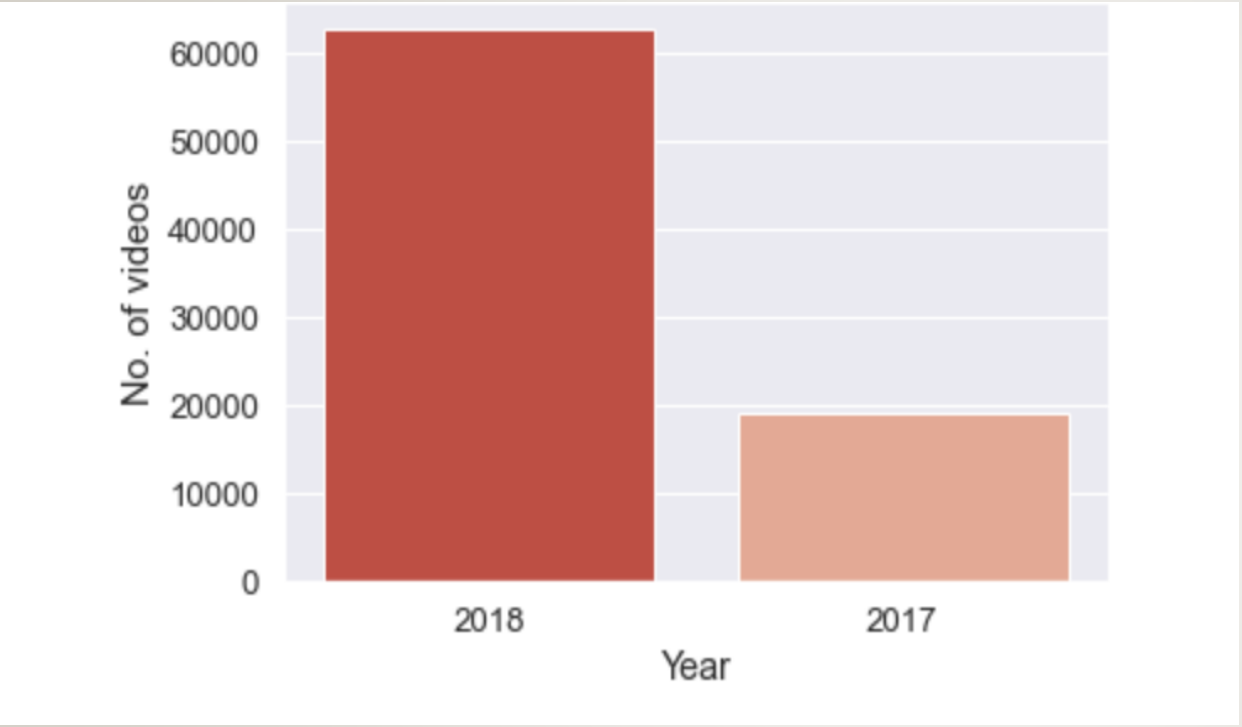
As we see the highest channels watched CNN, ESPN , The late show with Stephen.

SDAIA ACADEMY



plot shows videos that have 100,000,000 views and more have Title Length between 35 and 60 characters approximately.

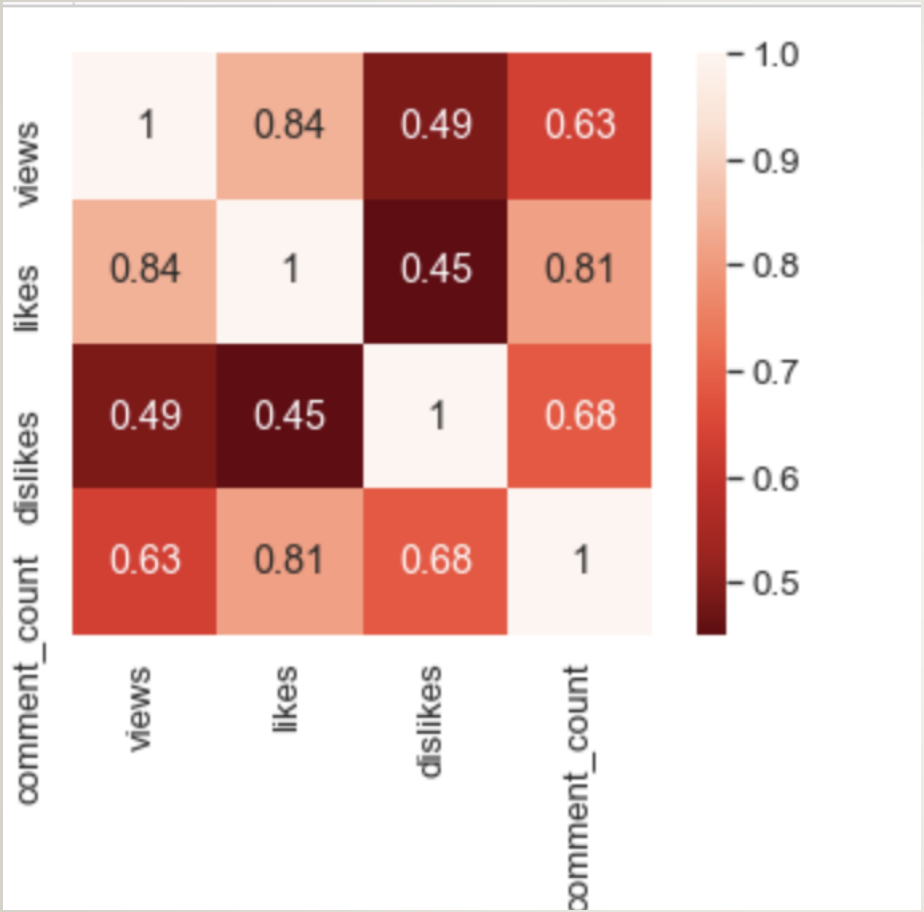
SDAIA ACADEMY



We can see that the dataset was collected in 2017 and 2018 with 77% of it in 2018 and 23% in 2017

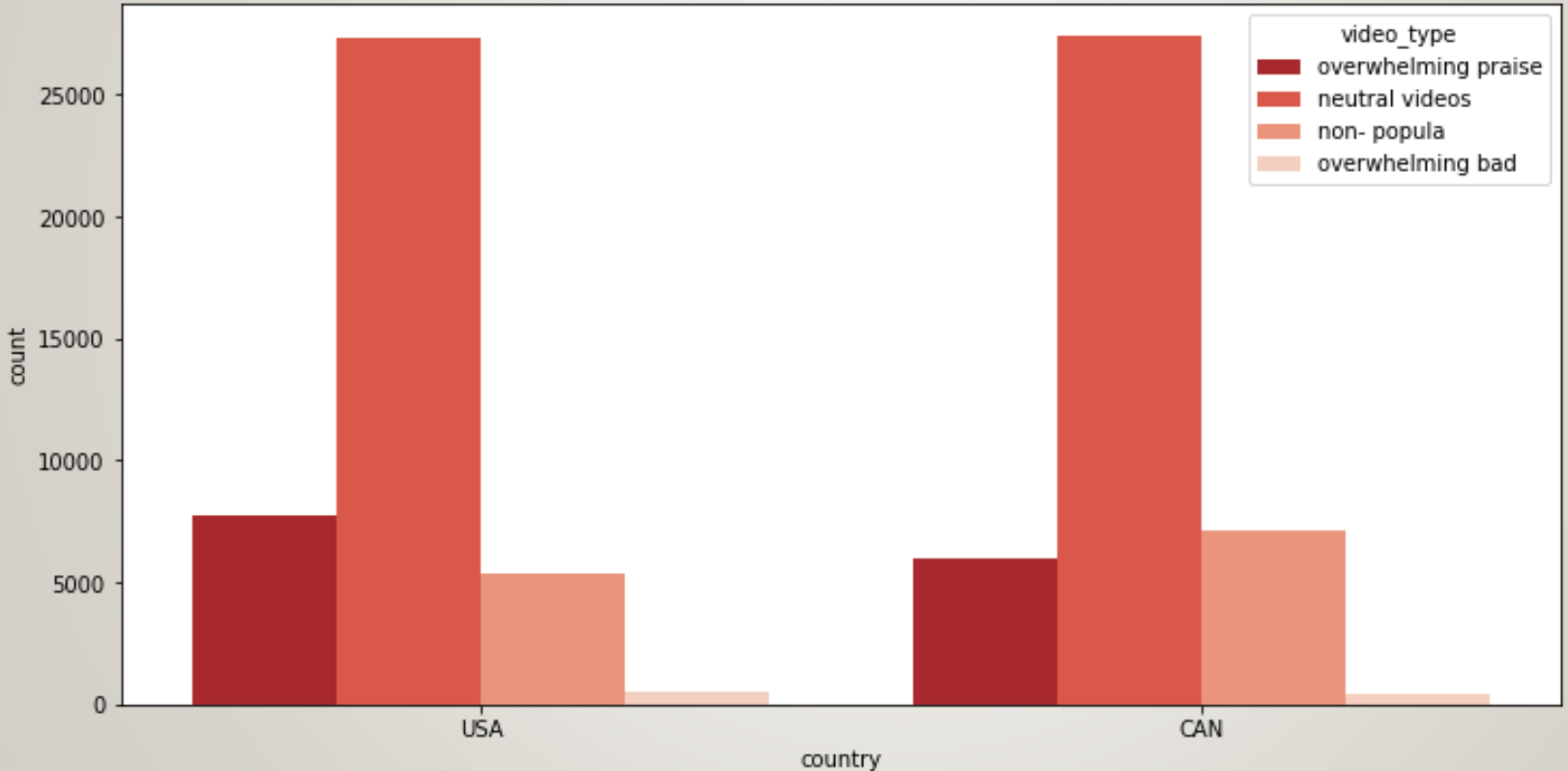
SDAIA ACADEMY

As we can see Trending videos contain words like Official, Trailer, Live, New, Challenge, HD etc. which helps attracting the viewer's attention.

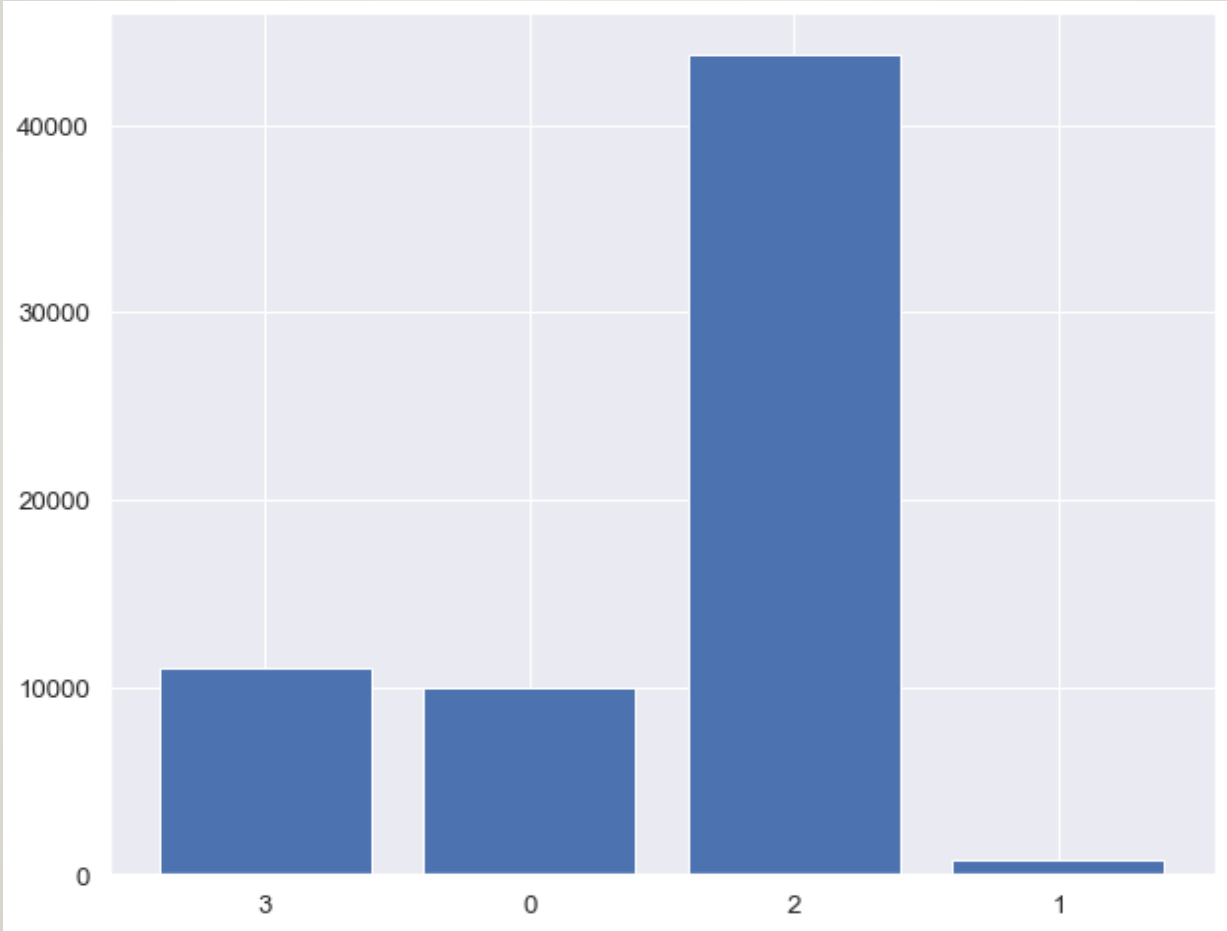


As we can see it is high correlation between Views and likes and likes and comment_count

As we see here the number of videos that become trending between the United States and Canada, the United States is higher in the emergence of trend



Imbalanced classes



FEATURES:

1. Views
2. Likes
3. Dislikes
4. Comment
5. Difference day between publish and trends
6. Tegs

MODELS RESULTS

Recall score	Model name
0.99	Random_forest
0.99	Voting
0.99	Bagging
0.94	knn
0.85	Decision_Tree
0.33	logistic

- **We can help a YouTuber to make their video as a trending one by telling them what day what time and what category of videos they should upload**

- **We can say if a video becomes a trending one when most of the people like that video**

It’s tell about audiences satisfaction

SDAIA ACADEMY

THANKS!

DO YOU HAVE ANY QUESTIONS?