Youtube Analytics

Big data final project

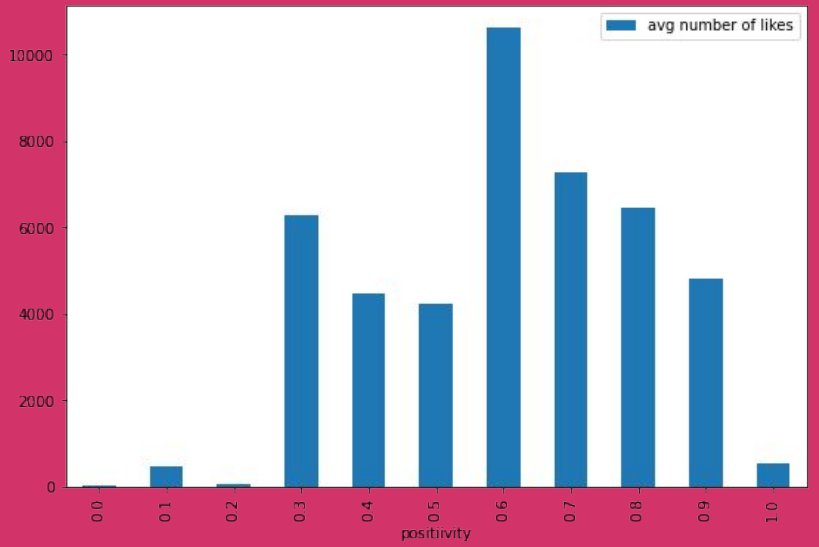
Sentiment analysis for youtube comments

- The level of engagement represented by the comments is much higher than by the reacts
- There is high correlation between discussion rank and the number of comments.
- The comments have more verification over the positivity or the negativity of the videos, compared to the reacts

Positivity

- Positivity is the ratio between positive comments and the total number of comments
- Visualization for the correlation between positivity and the average number of likes

Comment : There is some correlation between positivity and average number of likes , but we shouldn't ignore that there are many other parameters.



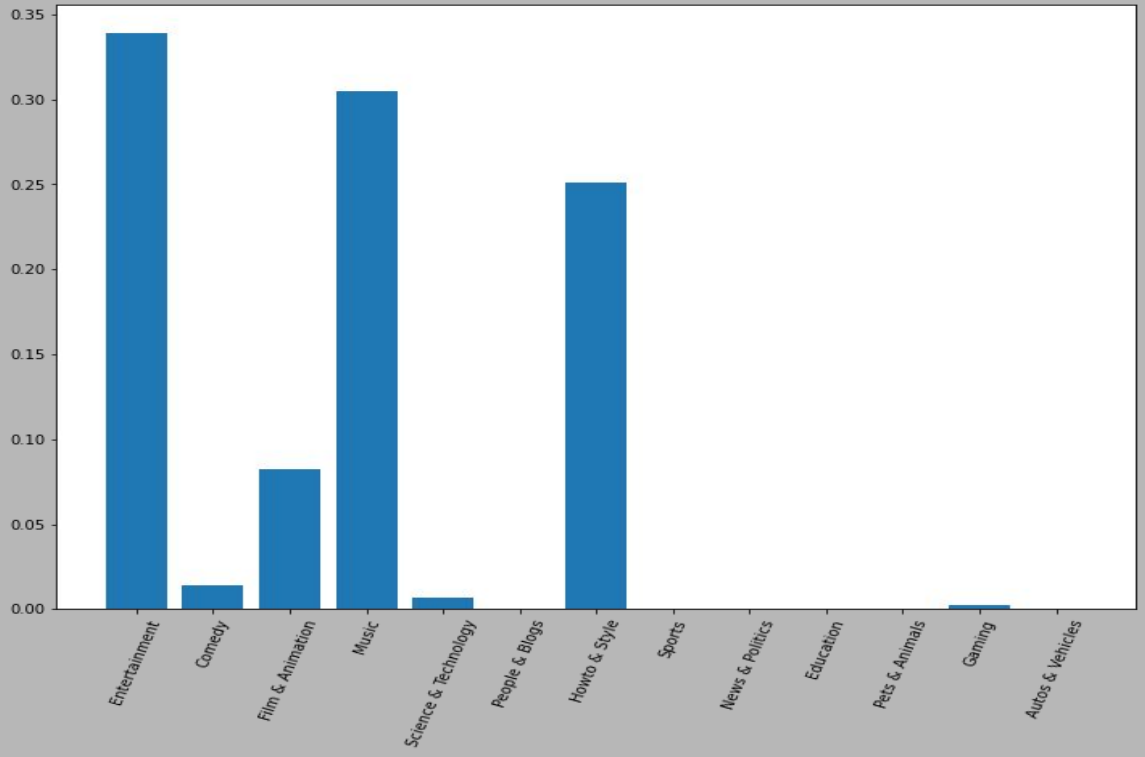
Category Distribution problem

- People ,definitely, has interest in more than one category through their experience in Youtube
- Video with one category doesn’t narrow down sufficiently for the individuals

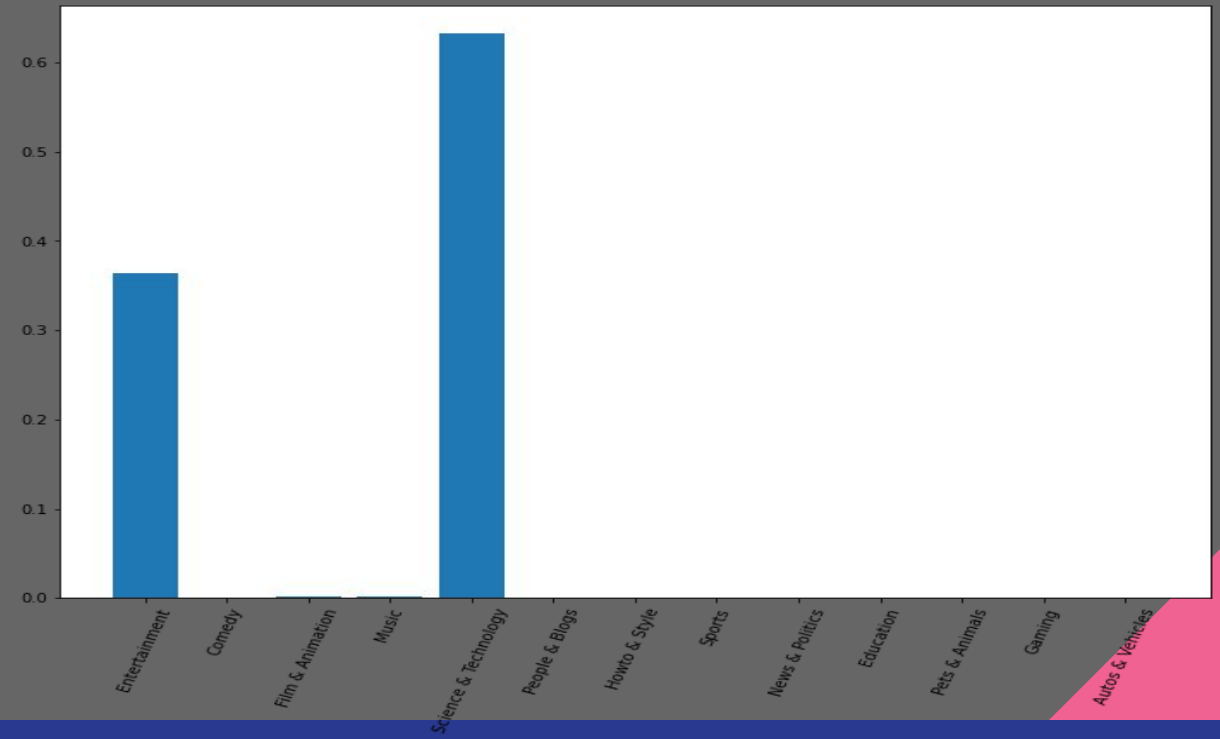
Category distribution

- We propose a machine learning model that map every video to its probability of its possible categories through its comments
- The model takes the comments of the video then predicts the probabilities of categories for the video
- Then, we can use this distribution to get the intersections of categories that this video belong to.
- That kind of feed to the recommendation engine will make it more efficient

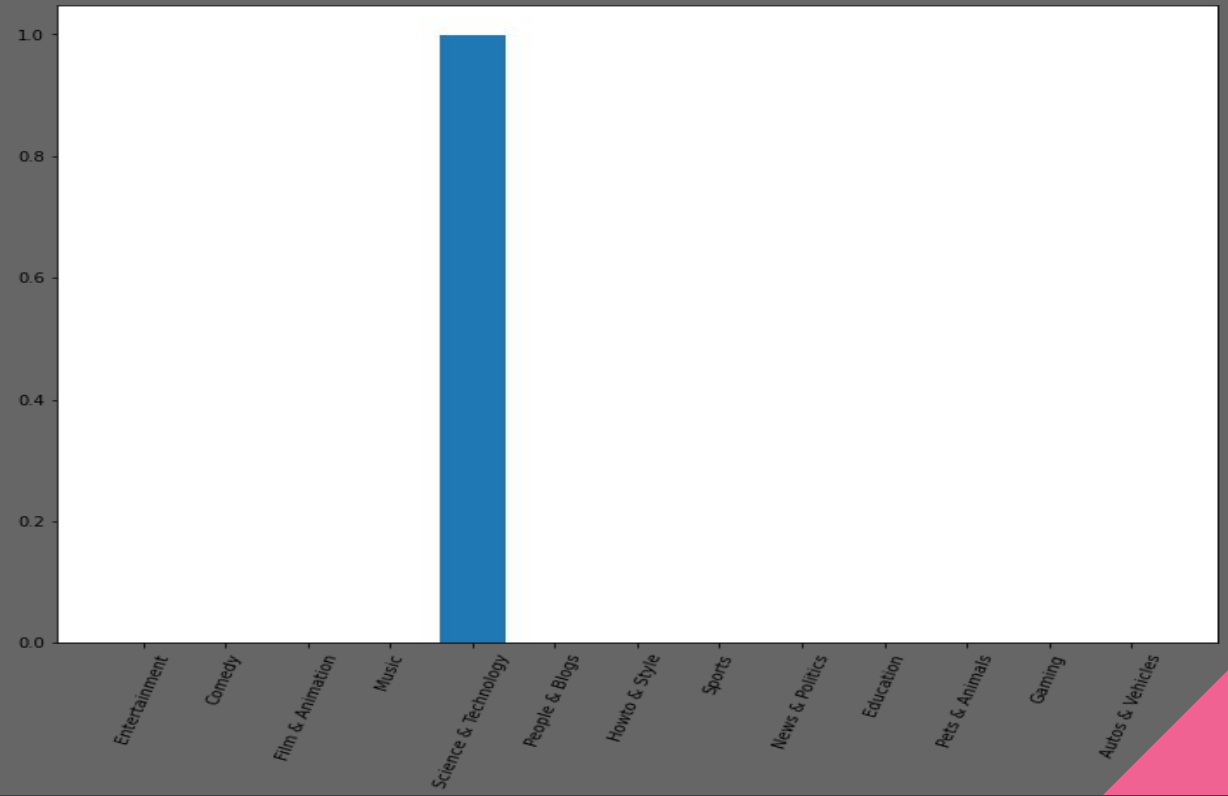
Comments category distribution



Comments category distribution

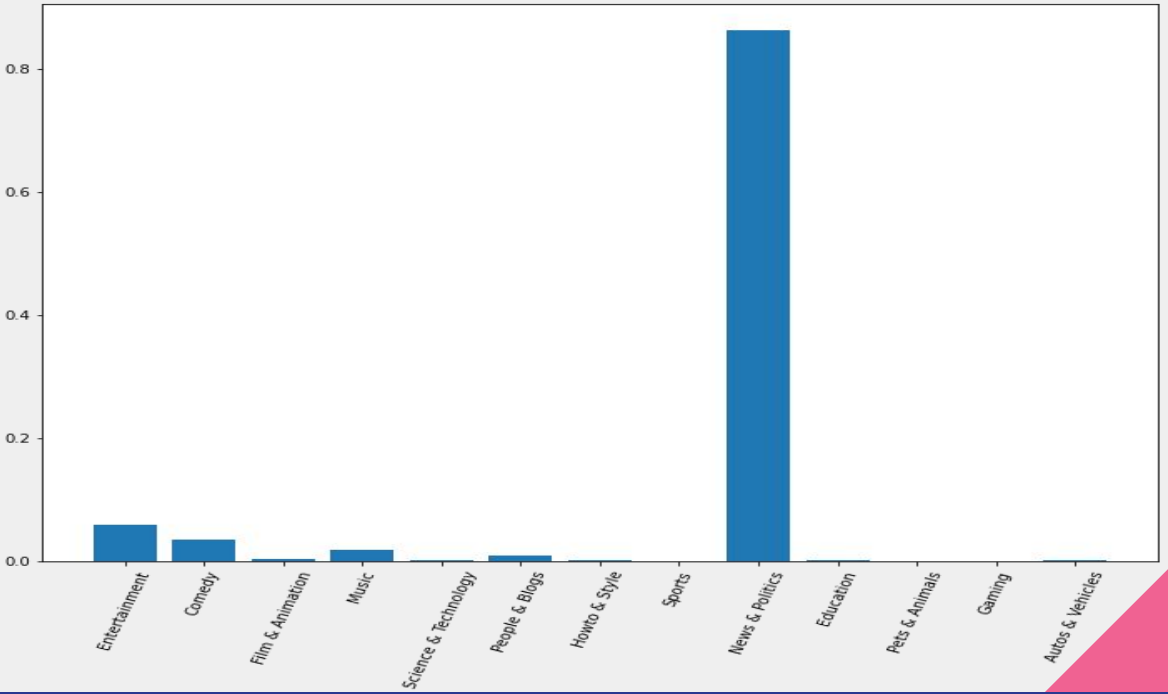


Comments category distribution



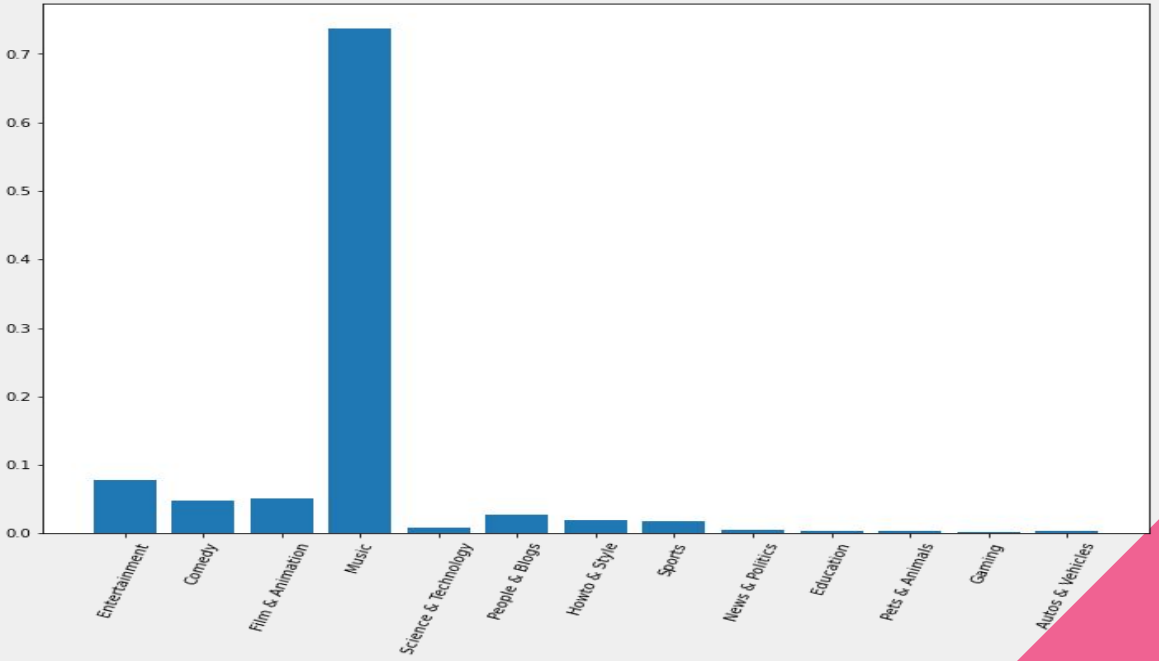
Video category distribution

Youtube link : https://www.youtube.com/watch?v=GGm0FQ6i74U&t=9s



Video category distribution

Youtube link : https://www.youtube.com/watch?v=eM_FR7I2Ttw



Video category distribution

Youtube link : https://www.youtube.com/watch?v=brfJuyTADuQ