
Predicting YouTube Video Views from Video and Thumbnail Features

Lean Ting Jin

Matrikelnummer 6956985
dummymail1@uni-tuebingen.de

Finn Springorum

Matrikelnummer 6124977
dummymail1@uni-tuebingen.de

Christian Traxler

Matrikelnummer 6969273
christian.traxler@student.uni-tuebingen.de

Anna Chechenina

Matrikelnummer 6987499
dummymail1@uni-tuebingen.de

Abstract

We are planning to use the YouTube API [1] to see which factors of a YouTube video best predict how many videos it will receive.

1 Introduction

2 Methods

Data Collection We collected 80k data points of YouTube videos from the YouTube API spread across 8 categories relating to entertainment (gaming, comedy, education, entertainment, how-to-style, news and politics, people and blogs, and sports) based on the pre-defined categories of the YouTube API. Additionally, the videos were queried to be relatively uniform from 2015 onward.

Video Statistics From the YouTube API, we saved relevant information such as the data of publication, view count, like count, comment count, subscriber count, duration, and the thumbnail of the video.

Thumbnail Image Statistics In addition, we computed some relevant thumbnail image statistics such as the hue (average hue in radians), saturation (average saturation normalized to $[0,1]$), lightness (average value/lightness normalized to $[0,1]$), contrast (standard deviation of the normalized grayscale image), and sharpness (variance of the Laplacian (second derivative) of the grayscale image).

Face Detection Finally, we use the DeepFace library [2, 3] to extract the number of human faces in the thumbnail.

Analysis

3 Results

4 Discussion/Limitations

Discussion

Limitations As a project based on the YouTube API, there are couple of inherent limitations of the data collection. One issue with the data collection is the categories: each video is either user-defined or automatically assigned only one category. This means that there exists some error in the categories such that errors are likely and that some categories that we wanted to exclude (like music videos) are likely to be included and may skew results.

Also,

5 Statement of Contributions

Here is an example:

XX performed the correlation analysis, organized the data and code for the processing of dataset1 and subdataset2, and created the scatter plot. YY created the random forest regression model, performed the data cleaning for the xyz analysis / xyz database, and created the bar charts to display the regression results. ZZ researched and collected the raw data, restructured the pipeline for the data analysis, and proof-read the draft for the final report. AA performed the data cleaning for dataset1, and performed the Ridge and Lasso regularization. All members of the group contributed to writing the report.

References

- [1] Google. YouTube Data API v3. <https://developers.google.com/youtube/v3/docs>.
- [2] Sefik Ilkin Serengil and Alper Ozpinar. A benchmark of facial recognition pipelines and co-usability performances of modules. *Bilisim Teknolojileri Dergisi*, 17(2):95–107, 2024.
- [3] Sefik Ilkin Serengil and Alper Ozpinar. Lightface: A hybrid deep face recognition framework. In *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pages 23–27. IEEE, 2020.