YouBait

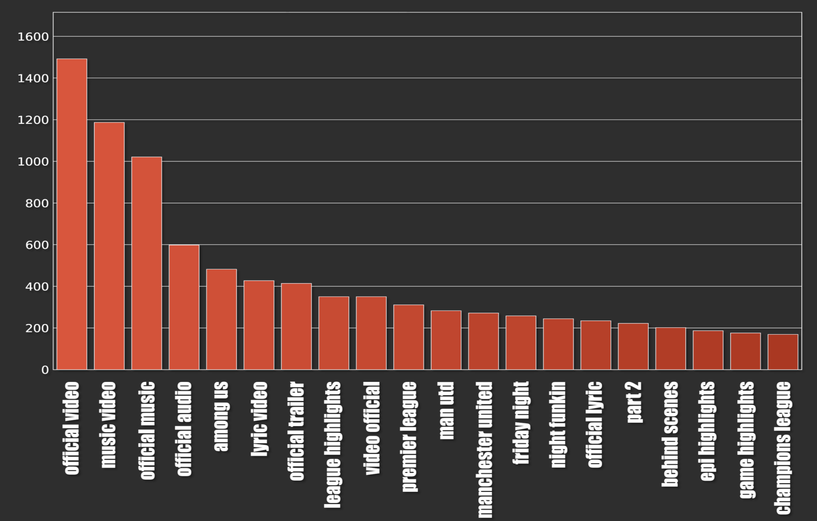
Search

DATA

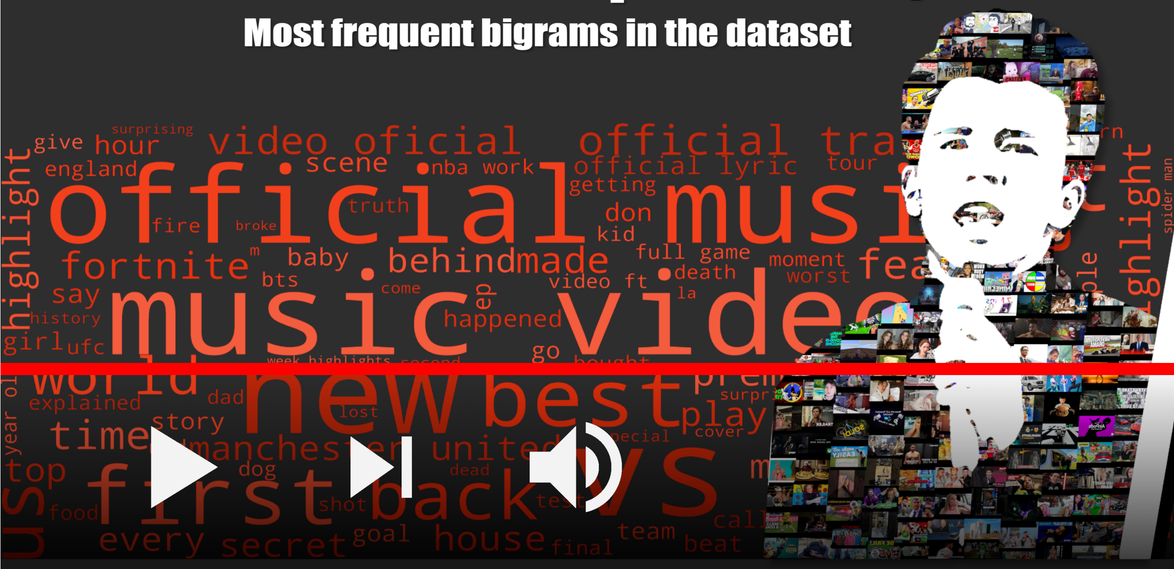
As a base for our dataset, we used a *YouTube Trending Video Dataset* from Kaggle. It contains information about **videos** in the ***Trending*** category (which potentially are more ***clickbaity***) and is updated daily.

The dataset, however, does not contain negative samples. To acquire those, we plugged into the *YouTube Data API v3* and downloaded data of all other videos from **channels** that appeared in the dataset.

We consciously chose to consider only a **thumbnail** and a **title**, as this is what a person usually sees when deciding whether or not to see a video.



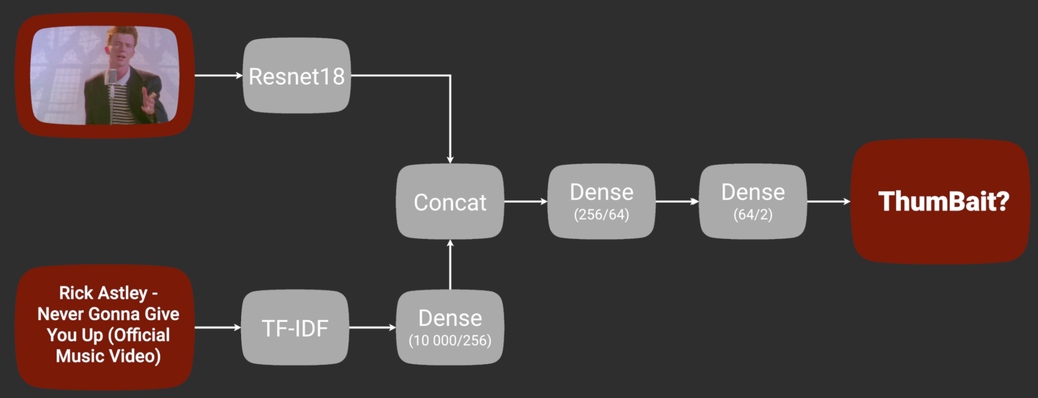
Most frequent bigrams in the dataset



We trained two **multimodal models** capable of processing both video thumbnails (images) and video titles (natural language). The first was responsible for predicting the **number of views**, the other - for classification whether a given video has the potential of becoming **trending**. The difference between the two is in the last, dense layer - classification expects the output to be of size *2*, regression - of *1*.

In our proposed model, two modules were used. For **natural language processing**, we selected TF-IDF capable of processing single words and bigrams. For **image processing**, we used the ResNet18 network. These modalities are subsequently combined into a single, video-specific representation.

To minimize overfitting, we used **dropout** layers between the fully-connected ones. Furthermore, we proved that replacing ResNet18 with **ResNet50** did not improve the results.



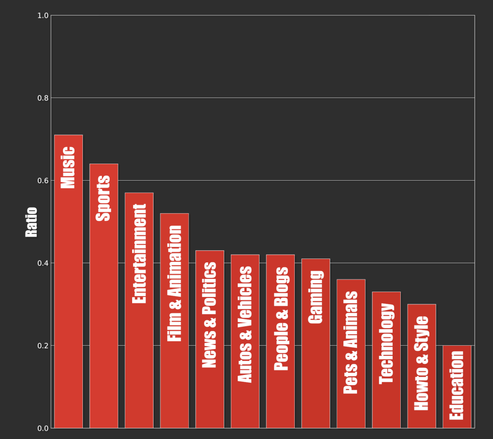
MODEL

RESULTS

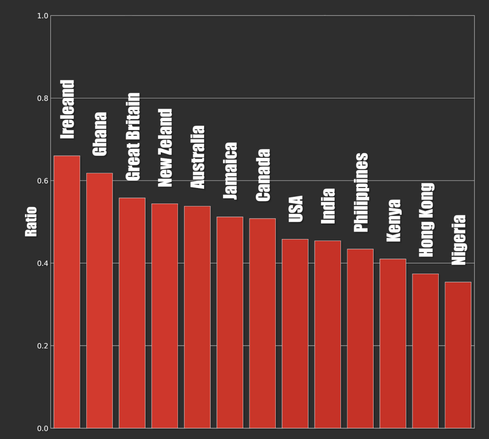
Notwithstanding the limited number of input features and the multitude of factors influencing a video becoming trending and its real number of views, the **classification** model achieves an **F1 score of 0.71**.

The **regression** model did not avoid this curse. It is worth noting that this task is much more complicated, both in terms of training the model, as well as evaluating it. We could not get **satisfactory** results within any metric; however, a clear correlation is noticeable.

We applied the trained classifier to a testing dataset to measure a ratio of clickbait videos considering their categories and countries they had been uploaded from. It appears that our model thinks **music** has the highest chance of becoming trending and it works best on data from **Ireland**.



Categories with highest clickbait ratio



Countries with highest clickbait ratio

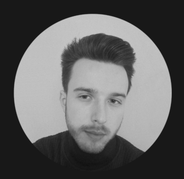
As a result of this project, an **interactive demo** has been created and, along with the **code**, can be accessed through QR codes.

Demo Code Bonus

Thumbait - is this project a bait?

HOW MANY?! views • 27 Jan 2022


Michał Karbownik


Michał Matusz


Patryk Szelewski

Wrocław University of Science and Technology

The aim of this project was to build a multimodal model, accepting images and text, that predicts whether a video appears in the *Trending* category on YouTube.

clickbait ***(noun)*** - (on the internet) content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page.

Thumbait, the name of our project, is composed of two words - *thumbnail* and *clickbait*. It represents the broad idea behind it, which is to predict if a given video is a clickbait based on its thumbnail image and title.