Question 1 - Exams 2

Ioannis Konstantakopoulos ,

st.numb:0960047

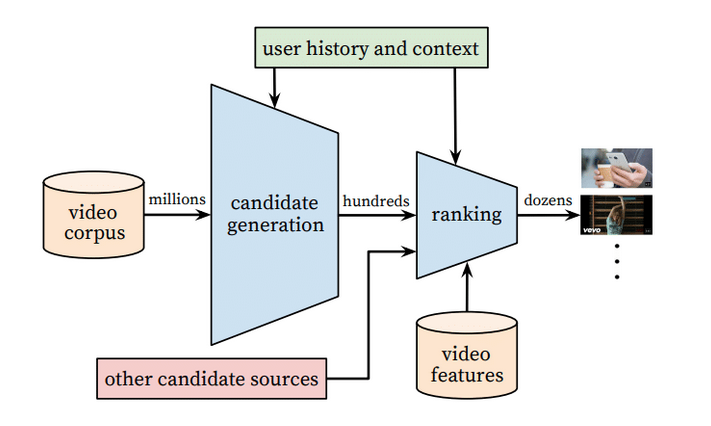
Answer

There are many videos on YouTube, both professional and amateur more than 500 hours of content are uploaded every minute. The amount of videos has to be narrowed down before thumbnails are even used. From them we can create some candidate generators which will window down the hundreds of videos that may be relevant to the user like the user’s search and watch history. The candidates from which we will generate the thumbnail will have the primary role of ranking the impression data to specialize and calibrate the predictions for the particular user interface.

Firstly, we have to explain what thumbnails are and what we should extract from video\_thumbnail. Typically they are well framed, in-focus, and focus on a specific subject (like the main character in the video). We can have this centered and detecting objects or persons from using COCO: Common Objects in Context(\*) (from our tutorial) . Meaning we can use deep neural networks(DNN) to identify if any non focus motion exists. Then we can compare the findings of the DNN and COCO and if the findings are in the same section then we can penalties the subject. By using COCO we can identify and centralize objects and a well framed parameter may be created. Furthermore thumbnails can contain a small text of 5-6 words to emphasize the video. To extract and identify the text we can use models like Tesseract (\*) . It’s important to contain no more than 9-11 words otherwise the score can be decreased. Additionally, thumbnails can create emotions, so we have to create thumbnails closer to the users interface in order to generate a higher chance to be clicked. Also will have higher rank in this feature.

Classifying videos as clickbait ensuring the quality for instance the words or the colors is a tactic that helps them capture more attention. Nevertheless, as (2) indicates “if you optimize for click-through rate you get clickbait, and if you optimize for watch time, you get incredibly long videos.”, so watch time has to be optimized otherwise it will be penalized. Finally, using a model like colorgram (\*) we can identify which colors are preferable and use them in the thumbnail, in some cases bright colors are preferred (landscapes) and in others darker ones get the preference (movies) and those should be rewarded accordingly.

To sum up, features we can construct from the video\_thumbnail column are like scores. The columns on nodes.csv might me like : i) score for focus ii) score for framing iii) score for clickbait (which will include the mean of the proportion of color score, watch time score and emotion score). We have to use the score accordingly to our investigation (user) and either favor them or not.



References

(1) <https://blog.hootsuite.com/how-the-youtube-algorithm-works/>

(2) <https://www.shopify.com/blog/youtube-algorithm#clickbait>

(3) <https://ai.googleblog.com/2015/10/improving-youtube-video-thumbnails-with.html>

(\*) All the \* are models from the manual I decided to use them as we were told to use only models from tutorial.