INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Software Development Group Project

Module no: 5COSC009C

2020/21

Predicting YouTube Video Popularity by Thumbnail

A Project Initiation Document by

Team Reboot Rebels

Sunera Munasinghe 20191270
Shehan Saleem 2018149
Samitha Kalpana 20191127
Nipun Shalinda 20191145
Udula Ratnayake 2017491
Praneeth Sankalana 20191024

October 2020

Table of Contents

List of Figures

List of Tables

1. Chapter 1: Introduction

YouTube is an online video-sharing platform which allows users to upload, view, rate, share, add to playlists, report, comment on videos, and subscribe to other users. YouTube is also the Second biggest search engine in the world. YouTube offers a wide variety of user-generated and corporate media videos. Available content includes video clips, TV show clips, music videos, short and documentary films, audio recordings, movie trailers, live streams, and other content such as video blogging, short original videos, and educational videos.

300 Hours of Video are Uploaded to YouTube every minute. 5 Billion Videos are Watched on YouTube every day. Compare that to the 800 Films that were Shown in American Cinemas in 2019, and to the 500 Scripted TV Shows that Were shown in American Television in 2019. These Massive numbers make the average user of YouTube who watch 160 Videos per day (A view is counted if at least 30 Seconds of the Video has been watched) statistically oblivious to the video content that is Actually uploaded on YouTube. That is the Reason that YouTube has an Algorithm, which they Claim can Connect the User to the Type of Content that they think they are interested in.

1.1 Project Scope

There are Multiple Factors that Contribute to the number of Vides a Video will get. Youtube.com Counts a View only if it has met the below 2 Conditions:

- The User has Watched the Video for 30 Seconds or More
- The User Clicked on the Video Voluntarily

First let us Discuss how a User will come across a YouTube Video.

- YouTube Home page Recommendation (This is likely if a user has watched one of your videos or a similar video recently.)
- YouTube Search Results
- A Video can Show up on Recommended Videos while a User is Watching a Video
- Subscribers that have Turned on Upload notifications will receive a notification that the channel has uploaded a video
- A Video is in a YouTube Playlist that shows up in Search Results.
- A Video can be embedded or shared somewhere else (Web page, Facebook, LinkedIn, Instagram, Twitter etc.).
- A Video can be featured as a trending Video in YouTubes Trending tab.
- Someone Directly Recommends a Video to their Friend/Family

In nearly each one of those scenarios, a User has the Ultimate choice of choosing to watch a Video or not to. “They will primarily use two criteria to decide if they want to watch your video or not. The first: does the title sound interesting, relevant, and like a topic that I would like to watch? And the second is the thumbnail: does it look relevant, enticing, clickable, and interesting?” (Landis-Eigsti, 2019)

The factors that Contribute to the number of Views a video gets are listed Below. This Video Includes Visual and Non-Visual Features.

1.2 Factors that Contribute to the Number of Views a YouTube Video

1.2.1 Thumbnail (Myers, 2020)

The Thumbnail of a YouTube video is the First thing users see when they Visit the YouTube Homepage. The YouTube thumbnail conveys most of the information a user will look for when deciding to watch a video.

“YouTube is a powerful way to market for your business. While Keyword optimizing your videos should be your #1 priority, designing your thumbnails is the 2nd most important way to get people to find and click on your video.” (Landis-Eigsti, 2019)

1.2.1.1 Faces

There has been Extensive Research done on how Faces attract Attention from the human eye. Multiple Research have been done to Prove and Challenge this Phenomenon.

Neil Patel says that “pictures of people are good. A page that has pictures of a person’s face encourages interaction and viewing and decreases a bounce rate. Use them as design elements on your website, on the about page and in social media profiles” (Patel, 2014, April 16)

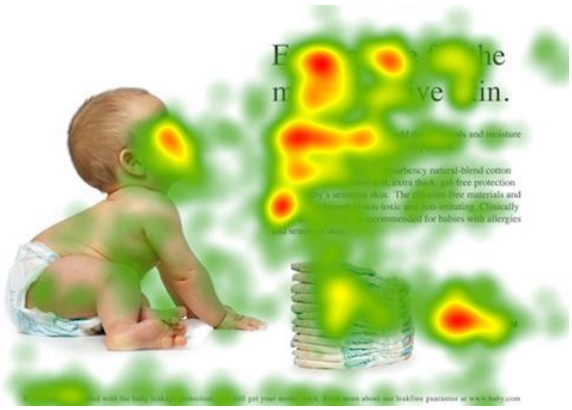


Figure 2: Eye tracker Heatmap (Patel. N 2014)

There are 2 Types of Faces that Play a Role on grabbing user attention, and these are:

- Human Faces
- Animals Faces
- Face of a Public Figure or Celebrity

1.2.1.2 Text

"Adding text headings on your video thumbnail holds many benefits, but the biggest benefit is that it gives the viewers more context about your video. A simple image, no matter how beautiful, isn't going to communicate that your video is a brilliant tutorial that will solve one of your viewers' most pressing problems." (Myers, 2020)

- Font size (Bigger is better)
- Font color
- Font type (Creative but readable) (Blade, n.d.)

1.2.1.3 Image Quality

Aditya Khosla, Atish Das Sarma and Raffay Hamid (Aditya Khosla, 2014) mention in their Research the different characteristics of an Image that may Predict its Popularity. This Includes Characteristics Such as Texture, Color Patches, Gradients etc. The Results Show the Correlation between the Popularity and its Characteristics for 20 Random Users. They also State that Social Cues play a part in Predicting Popularity of an Image.

1.2.1.4 Design Elements

Design Elements Such as Icons, Large Punctuations could influence the Number of Clicks the Thumbnail Gets.

1.2.1.5 Consistent Style

Having a Profitable Design Style for Video thumbnails with Consistent Branding is something multiple YouTube Creators do to Maintain their Brand Image with their Audience.

1.2.1.6 Similarity between title and thumbnail image content

YouTube Counts a View if the User Watches the Video for more than 30 Seconds in Total. If the User Discovers that the Thumbnail and the Video has nothing Common the User may not Continue watching that Video.

1.2.1.7 NSFW Score

Although The YouTube Custom Thumbnail Polity does not Allow any form of "pornographic images depicting sexual acts, genitals, or fetishes for the purpose of sexual gratification" (Google, n.d.) YouTube video Creators find ways to Include Cleavage baring bodies as thumbnails to drive traffic as discovered by (Butcher,2018). Where he analyzed the YouTube Reply Girls, where a group of women

"used their cleavage baring bodies as thumbnails to drive traffic" to their Videos between 2011 and 2012 (Butcher, 2018, p. 128).

1.2.1.8 Clickbait Score

Keywords and Phrases such as "Unbelievable" and "You Would not Believe" and Multiple types of Keywords and Images that some YouTube Creators use to maximize the number of Clicks they get has been a Phenomenon that has been researched extensively. There has been research done to Predict and Hide these links as well.

Clickbait is Simply a Teaser, that sometimes may be deceptive to make a user click on a video and may leave a user dissatisfied. "However, the teaser information alone does not suffice to detect clickbait videos" (Jiani Qu, 2018)

1.2.2 Date of Uploading

Users may Look at the Date of Uploading of a Video before Clicking on it based on the Purpose of Browsing YouTube

1.2.3 Title of the Video

The Title of the Video is The Other Major Component that Affects if a User will Click on a Video

1.2.4 Number of Views

The Need for Social Proof is Validated by the Number of Views a Video already has. A User may tend to Click a Video that is More Popular Over a Video that has not been Watched yet.

1.2.5 Length of Video

The User may look at how long a Video before Clicking on it. Educational Videos may do better if they are longer, and Entertainment videos may do better if their Shorter.

1.2.6 Niche of the Video (Educational/Entertainment/Music etc.)

Based on the Primary reason a user uses YouTube the niche of the video based on the users' interest may affect the CTR

1.2.7 Channel Name and Image

The Readability of the Channel name and the Quality of the Image may Influence the User to Click on the Video

1.2.7.1 Familiarity of the Person to the audience

The Familiarity of the Channel and the Verified Status may affect the CTR

1.2.8 Relevancy to the current world events

The Relevance of the Video Title and Thumbnail with the Most Popular Current World events can Affect the CTR and View Count.

1.2.9 Other Factors

Multiple Factors that have not been stated such as the Age of the User, the Purpose of the User, the Internet Access Level of the User could also impact the CTR.

1.2.9.1 Cultural

Cultural Factors Such as the Language of the User will affect if the user is interested or not.

1.2.10 Likes and dislike ratio

A Video with more Dislikes may Influence a User to Leave the Video before watching 30 Seconds of it to be Counted as a View.

1.2.11 Quality of the Video Itself

The Quality of the Video may affect the Number of Views a Video will get as a Video needs to be able to keep attention to the user for at least 30 seconds to be counted as a view.

1.2.11.1 Score of Keeping the Users Attention

Score of Keeping the Users Attention for 30 Seconds (You need a Minimum of 30 Seconds Watch time to be considered as a Video)

1.2.11.2 NSFW Score of the Video

The NSFW Score Within the Video may keep or Lose the Users Attention.

1.2.11.3 Audio Quality

1.2.11.3.1 Well Mixed Audio

Audio to Hear the Words Spoken etc. are Important to keep the Attention of a User.

1.2.11.4 Video Quality

1.2.11.4.1 Cinematography

The Stability of the Movements of the Video etc. Can also Affect the View Count of the Video

1.3 In Scope

In This Project we will be Ignoring the Factors Such as: Channel Name, Video Title etc. and We will only be Doing our Research on the Thumbnail. This Research is Limited to Entertainment Videos by Small Individual/Independent YouTube Creators.

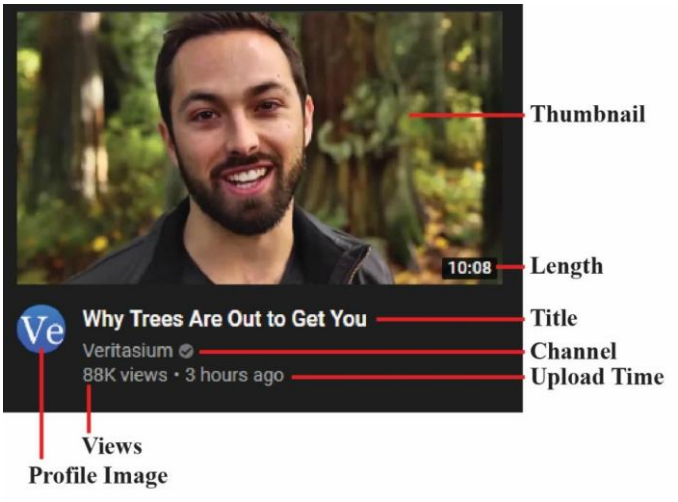


Figure 3: Visual Properties that a User will Use to Decide to Click on the Video

Giving a Score for user uploaded Thumbnails based on overall appearance by analyzing key components as mentioned in the above section (Factors that contribute to click rate).

This Research will only be Focused on the Thumbnail of a Video and how the Following aspects affect the number of Views a Video will get along with

1. Thumbnail
    a. Faces
        i. Human/Animal
        ii. Face of a Public Figure
    b. Text (Font etc.)
    c. Image Quality
    d. Design Elements/ Icons etc.
    e. Consistent Style
    f. Bright/dark theme according to the content mood
    g. Similarity between title and thumbnail image content
    h. NSFW Score
    i. Clickbait Score

1.4 Out Scope

This Research has unlimited potential to Help YouTube creators to run their Video Through before uploading it to the Web. Some of the Features Could be: Analyzing the Video to estimate Watch time based on all the Possible Variables will be the most beneficial for a YouTube Creator. Although Google's DeepMind algorithm is working on Developing an Artificial Intelligence that can accurately predict these factors, there is no Software for YouTube Creator to use to Evaluate their Video Uploading.

This Research will not Include the Following Factors that affect the View count of a YouTube Video:

- Date of Uploading
- Title of the Video
- Number of Views
- Length of Video
- Niche of the Video (Educational/Entertainment/Music etc.)
- Channel Name and Image
- Relevancy to the current world events
- Other Factors
- Likes and dislike ratio
- Quality of the Video Itself

The final solution will only be predicting the popularity considering only the thumbnail of the video. The Dataset that will be used to Train the Algorithm is Limited to Independent YouTube Creators. That have At least an Average of 10,000 Views Per Video

1.5 Problem Domain

1.5.1 Problem Background

When someone uploads a Video to YouTube, The Video Creator has a choice to set a Thumbnail to the Video. If the Video Creator does not choose that Option, then YouTube will Give the Creator a Choice from 3 Frames of his Video for the Thumbnail. (Szabo & Bernardo, 2010)

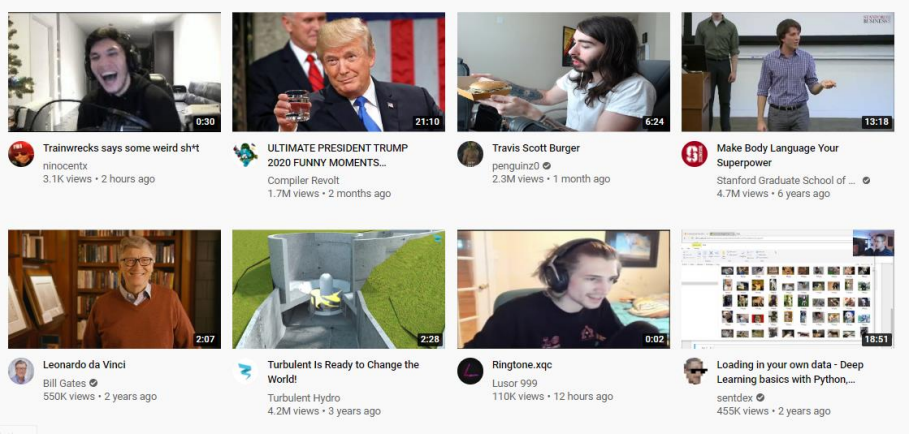
Most Regular Video Creators Choose to Create their Own Thumbnail because it gives them a higher ranking in SEO. YouTube has also introduced a tool called “Real time CTR” for YouTube creators. (Borghol, et al., 2011) This helps these Video Creators to go an extra Step further to Create Multiple Thumbnails. Within the First 30 Minutes of Uploading a Video the Youtuber has the ability to Quickly Change the Thumbnail several times to check which thumbnail gets the highest Click Through Rate. the most Views on the Video. The Click Through Rate has a Massive impact on the Success of a YouTube Video. (Bandari, et al., 2012), (Pinto, et al., 2013) (Khosla, et al., 2014) (Xu, et al., 2014)

CTR = Number of Clicks on Video/Number of Impressions

One could argue that the Objective of the General YouTube Video Creator is to Maximize the Profit they Earn while having the satisfaction of Creating Original and Content that is aligned with their Passions or Expertise. While the Objective of Alphabet (The Mother organization of YouTube) is to Maximize the Long-Term Profit that they Earn.

1.5.2 Research Gap

Several Attempts have been made to Predict Video Popularity. There also Have been Attempts to Automatically Select a Thumbnails from the Frame of a Video. (Song, et al., 2016)



From Our Findings, there is no Software that Gives a Score to Thumbnails Uploaded by a User.

Some Features of Existing Software's are:

- Choosing the best Frame from a Video for a Thumbnail
- Clickbait Detection,
- Predicting Popularity of Online Videos etc.

This Solution will use Machine Learning to Learn what type of Thumbnail will get them the Most Views.

1.6 Feature Comparison Chart

Commented [1]: Proposition to change "Product 1" to "Our Product"

Feature	Product 1	Product 2	Product 3
Uploading User Created Thumbnails	✓	✗	✗
Giving a Score for user uploaded Thumbnails	✓	✗	✗
Ability to Recommend Improvements to be Made to the Thumbnails	✓	✗	✗
Visualizing the Predicted Popularity of the Video	✗	✓	✗
Identify the regions of the image with imprinted subtitles.	✗	✓	✗
Suggesting a new thumbnail by generating from the video	✗	✗	✓

Table 1 - Feature Comparison Chart

Product 1 – The Final Solution of this research project

Product 2 - Predicting popularity of online videos using Support Vector Regression (Trzcinski & Rokita, n.d.)

Product 3 - Selecting YouTube Video Thumbnails via Convolutional Neural Networks (Arthurs, et al., n.d.)

1.7 Features of the Prototype

Some of the main features of the prototype will be mentioned below.

1 Ability to Upload up to 3 User Designed Thumbnails
2 Retrieve a Score for Each Thumbnail out of 100 Based on its Predicted CTR
3 Recommendations on how to Improve the Thumbnails (Such as: Improving Image Quality, Including a Face in the Thumbnail etc.)

2. Chapter 2: Literature Review

2.1 Introduction

In the light of the recent years, social media platforms have started to gain massive popularity among the public. Since social media platforms are among the fastest growing platforms in the software world, more and more child software is needed to help users with their usual day to do tasks related to the relevant social media. Among those platforms, YouTube is one of the biggest platforms available, with more than 5 billion daily video views, and with an overall total of 1,300,000,000 daily users (merchdope.com, 2020). Nearly 300 hours of videos are uploaded every minute to YouTube too. With such a massive creator base for the video streaming site, it's almost crucial to have a thumbnail scoring system to let the user know whether the custom thumbnail that he/she has created would rack up clicks for the video.

2.2 Review Contents

This literature review aims to cover the initial in scope and the out scope of the project. Furthermore, it aims to cover out various other research conducted by other researchers, whom will be cited in the review and on the **References** too. The feature comparison chart will be available in its relevant topic section too, and it will be given a in depth view too

2.3 Literature review of the Domain

2.3.1 Introduction to Thumbnail Scoring Model

A thumbnail is the main deciding factor which chooses whether a viewer clicks on a video or not. A thumbnail has several factors which helps it stand out from the other thumbnails displayed on the page (Refer **Factors that Contribute to the Number of Views of a YouTube Video**). Among them, some crucial factors are Design Elements, Similarity between thumbnail and video. An application that could analyze the image elements of the thumbnail could use machine learning capabilities that could help it to recognize various factors of the thumbnail and rate the thumbnail based off the suitability of the images. The application would have an option to let the user upload a thumbnail to the application, a feature not yet seen in other software. The thumbnail would be analyzed using the image detection algorithm and scored, with regard to more than 10 factors.

2.3.2 Existing Research/Software

Some research has already been made on the subject and published to the internet. One such is the research done on “Automatic Selection of Thumbnails from Videos” (Yale, Miriam, Jordi, Alejandro, 2016).

Commented [2]: Our model unique to us

2.3.2.1 Automatic Selection of Thumbnails from Videos (Yale, Miriam, Jordi, Alejandro, 2016)

The research shows that the aforementioned software uses 3 key steps in the process of selecting a suitable and relatable thumbnail. The 3 main steps are Frame Filtering, Keyframe extraction and Thumbnail Selection. Frame filtering refers to the process of filtering out unwanted frames from the video to have a general set of frames to be used in the Keyframe extraction process. The techniques used in the frame filtering process have the ability to filter out blurry, dark, and uniform frames. It leaves the application with colorful, still, and sharp images to choose from. By the Frame Filtering technique, 88.64% of frames could be wiped off on average.

The second step is to extract a keyframe from a scene. The primary objective of this step is to discard duplicate frames. This process is much more forgiving than the prior since it is better to over-saturate the frame set rather than under-saturate it. It uses “” Clustering Analysis” to filter out unwanted frames by selecting a cluster and selecting a frame on the centroid of the scene on the cluster. It was also highlighted that the stillness of the certain frame was a significant factor when picking the keyframe.

The third and the final step to select a thumbnail is the thumbnail selection itself. The application analyzes the color, texture, quality, and composition of the remaining frames. It is all up to the aesthetics of the frame. After all the frames are aesthetically analyzed, the final remaining frame would be chosen after some lengthy mathematical calculations.

It should be noted that the application only selects the thumbnail from the existing video. It does not engage in any modification whatsoever. Furthermore, it does not give the user the freedom to input a frame which he/she thinks is attractive.

3. Chapter 3 System Requirements Specification

3.1 Purpose

The purpose of this document is to present a detailed description of the video popularity prediction solution. It will explain the purpose and features of the system, the interfaces of the system, what system will do, the constraints under which s must operate and how the system will react to external stimuli.

3.2 Intended Audience

This document is intended for both the stakeholders and the developers of the system.

3.3 Product Scope

Mainly there are two main factors related in order to count as a view of a YouTube video. Those 2 are, the user should watch the video at least 30. The user clicked on the video voluntarily.

There are so many reasons for users to implement the above factors. But this project is forces for only thumbnails. The thumbnail is the first thing the user see when the user visit in home page of YouTube. So, the thumbnail conveys most information the user will look for when decide to watch the video. In order to get the user attraction, the following factors are considered in this project

1. Faces
    - Human faces
    - Animal faces
    - Face of a public figure or celebrity
2. Text
    - Font size
    - Font color
    - Font type
3. Image quality
    - Texture
    - Color patches
    - Gradient

3.4 Overall Description

3.4.1 Product Perspective

3.4.2 Product Functions

3.4.3 User Classes and Characteristics

3.4.4 Operating Environment

3.4.5 Design and Implementation Constraints

3.4.6 User Documentation

3.4.7 Assumptions and Dependencies

3.5 External Interface Requirements

3.5.1 User Interfaces

3.5.2 Hardware Interfaces

3.5.3 Software Interfaces

3.5.4 Communications Interfaces

3.6 System Features

3.6.1 System Feature 1

3.6.2 System Feature 2

3.6.3 System Feature 3

3.7 Other Nonfunctional Requirements

3.7.1 Performance Requirements

3.7.2 Safety Requirements

3.7.3 Security Requirements

3.7.4 Software Quality Attributes

3.7.5 Business Rules

3.8 Other Requirements

4. References

Aditya Khosla, A. D. S. R. H., 2014. *What makes an image popular?* [Online]
Available at: https://people.csail.mit.edu/khosla/papers/www2014_khosla.pdf
[Accessed 23 10 2020].

Arthurs, et al., n.d. *Selecting YouTube Video Thumbnails via Convolutional Neural Networks*, s.l.: s.n.

Bandari, R., Asur, S. & Huberman, B. A., 2012. *The Pulse of News in Social Media: Forecasting Popularity*, s.l.: s.n.

Bishop, S., 2020. Algorithmic Experts: Selling Algorithmic Lore on YouTube. *Platformization of Cultural Production*, 6(1).

Blade, S., n.d. *Social Blade.* [Online]
Available at: https://socialblade.com/youtube/education/101/10/designing-custom-thumbnails-for-youtube
[Accessed 11 10 2020].

Borghol, Y. et al., 2011. *Characterizing and modelling popularity of user-generated videos*, New South Wales: University of New South Wales.

Butcher, 2018. Cleavage-control: Stories of algorithmic culture and power in the case of the YouTube "Reply Girls." In Z. Papacharissi. In: *A networked self and platforms, stories, connections.* s.l.: Routledge, pp. 141-159.

Butcher, T., 2016. Neither black nor box: Ways of knowing algorithms. In S. Kubitschko & A. Kaun (Eds.). In: *Innovative methods in media and communication research.* s.l.: Palgrave Macmillan, pp. 81-98.

Google, n.d. *YouTube Custom thumbnails policy.* [Online]
Available at: https://support.google.com/youtube/answer/9229980?hl=en&ref_topic=9282679
[Accessed 24 October 2020].

Jiani Qu, A. M. G. P., 2018. Towards Crowdsourcing Clickbait Labels for YouTube Videos. *Bauhaus-Universität Weimar and Leipzig University*.

Khosla, A., Sarma, A. D. & Hamid, R., 2014. *What makes an image popular in Proceedings of International World Wide Web Conference.* s.l., s.n.

Landis-Eigsti, J., 2019. *Why Choosing the Right Thumbnail on YouTube Is So Important?* [Online]
Available at: https://jacoblevideoproduction.com/why-choosing-the-right-thumbnail-on-youtube-is-

so-important/
[Accessed 23 October 2020].

Myers, L., 2020. *Louise Myers Visual Social Media.* [Online]
Available at: https://louisem.com/198803/how-to-youtube-thumbnails
[Accessed 20 10 2020].

Patel, N., 2014, April 16. *8 Powerful Takeaways from Eye Tracking Studies.* [Online]
Available at: https://www.quicksprout.com/eye-tracking-studies/
[Accessed 23 October 2020].

Pinto, H., Almeida, J. M. & Gonc¸alves, M. A., 2013. *Using early view patterns to predict the popularity of YouTube videos*, s.l.: s.n.

Song, Y., Redi, M., Vallmitjana, J. & Jaimes, A., 2016. *To Click or Not to Click: Automatic Selection of Beautiful Thumbnails from Videos*, s.l.: s.n.

Szabo, G. & Bernardo, A. H., 2010. Predicting the Popularity of Online Content. p. 9.

Trzcinski, T. & Rokita, P., n.d. *Predicting Popularity of Online Videos Using Support Vector Regression*, s.l.: s.n.

Xu, J., van der Schaar, M., Liu, J. & Li, H., 2014. *Forecasting popularity of videos using social media*, s.l.: s.n.