CS 6474 - Social Computing Project Report

YouTube Videos Folk Theories, Too Good to be True?

1. Introduction

YouTube is a video sharing platform, which is used by a wide range of content creators. There is increasing interest amongst these video creators to find ways to optimize the reach of their video based on attributes such as the usage of tags. However, the algorithms YouTube uses to curate, select and present information are opaque to users. As a result, folk theories about how these curation algorithms work have been developed. In this project, we aim to test several folk theories that try to explain why YouTube videos go viral. A folk theory is an idea or assumption developed by ordinary people which is used to predict or explain a certain behavior, and does not involve a high level of technical knowledge.

2. Significance

The study of virality, in general, is of great importance since it allows us to glimpse at the ways in which the online and offline world affect one another. For example, it is widely believed that the 2008 presidential elections in the USA were significantly influenced by the virality of the "Yes we can" video of Obama on YouTube. The content that goes viral on any social computing platform allows us to have a better understanding of prevailing issues concerning people offline. Additionally, what goes viral on social computing systems could shape interactions within society.

There are several reasons why certain posts/videos go viral on social computing systems and individuals are constantly trying to figure out the "secrets" to virality. In this study, we aim to test folk theories related to viral videos on YouTube by evaluating the reliability of the metrics mentioned by them. We would do so, by looking at whether the viral videos on YouTube exhibit the characteristics that the folk theories claim they should have. By testing the different folk theories present, we can evaluate how well the folk theories approximate the YouTube's curation algorithm.

3. Related Work

User-Generated Content (UGC) has been the driving force behind popular online social media platforms, especially video content systems, like YouTube, that almost entirely rely on it. In "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System," [1] Cha et al. (2007) study the impact of YouTube and the different properties associated with the content of its videos. Through a data-driven analysis, they find that a peer-to-peer (P2P) communication system is imperative for massive content distribution and describe how the YouTube structural framework could benefit from it. Since the publishing of this work, YouTube has made improvements regarding P2P communication which caused content to be shared more rapidly. This study was extensive and focused on several aspects of YouTube and other UGC video systems such as user behavior, interface design, and content aliasing.

In 'What makes a video go viral? An analysis of emotional contagion and Internet memes' [2], the authors, Rosanna E. Guadagno et al. explain the various factors that cause a video to go viral. They especially touch on the various emotional and psychological factors that cause people to forward certain types of videos over others. The authors' analysis methodology is through a survey. The authors use their self-assigned categories, such as 'cute', 'humorous' along with other factors such as emotion associated with the videos to measure the likelihood of forwarding the video for participants. They also go into theorizing the impact of factors such as social validation, the source of the video to determine why participants chose to forward certain video content over others. One of the major limitations that the authors' acknowledge is that the study uses a culturally homogeneous sample of participants and generalizes the results from the survey to other population. Also, the age range of the participants was also an issue: the participants in this study were undergraduates. The study is limited because of the methodology of using surveys. In contrast, the data we've collected will not be constricted by age or culture.

In ‘Viral Actions: Predicting Video View Counts Using Synchronous Sharing Behaviors’ [5], David A. Shamma, et al. aim to predict the number of likes that a YouTube video might gain. They use the data from the sharing data of usage from the video-sharing environment: Zync (a plugin for yahoo messenger app), a platform that allows users to view and interact with a video simultaneously during a chat session. The authors’ argue that implicit sharing in a more accurate contributory factor to the possibility of likes that a video might get. The basic limitation of this paper is that it is basing the chat platform as the major source of video sharing that takes place. Contrary to the assumption we see that YouTube videos are mostly shared on Twitter and Facebook. Our data sources allow us to bypass such assumption by directly accessing the view count, rather than aggregating from multiple sources of sharing.

In “Virality over YouTube: an empirical analysis” [7], Khan and Vong look at the determinants of a YouTube video’s virality, by analysing the “social” (eg: number of subscribers) and “non-social capital” (eg: duration of the video) of YouTube videos and their creators. Amongst their several observations is the fact that music videos are disproportionately likely to go viral on YouTube. However, their analysis was largely restricted to english language videos and the dataset only contains 100 videos. Music has been believed to add more quality to viral YouTube videos. In our study, we will be testing this folk theory.

In “First I ‘like’ it, then I hide it: Folk Theories of Social Feeds” [25], Motahhare Eslami et al. curate a list of folk theories that seek to explain Facebook’s News Feed curation algorithm from interviewing 40 diverse participants who were users of the social media platform. For participants who were initially aware of the curation algorithm, their folk theories included believing that “[t]he more people that click that they like [a story], the more people that comment, the more people get to see it.” The researchers found that by exposing the unaware participants to Facebook content they missed because of the curation algorithm, they developed similar folk theories as the aware participates. Our project will curate folk theories related to how YouTube’s curation algorithms work and what characteristics increases the chances of a video becoming viral.

A number of blogs exist on the internet that shares various strategies to make viral videos on YouTube. The most common strategies are - (1) Viral videos generally have an intriguing and catchy title [3]. (2) Keywords and tags go a long way in determining the category of the video and help in reaching the target audience. (3) Viral videos are generally candid, fun, entertaining and almost always convey some message [4]. This message may be direct or embedded in the video subtly. (4) Viral videos are a team effort [5]. Many individuals are involved in the background who contribute to the success of the video. (5) The platform in which the video is published makes a difference. (6) Once a video gets traction, it is likely to spread like wildfire and viewership grows exponentially - commonly known as the snowball effect. (7) Researchers have concluded that viral videos are produced by only a few YouTube channels [6]. We will consider each of the characteristics mentioned while determining what folk theories will tested.

4. Objectives, Goals and Outcomes

In this study we aim to test the efficacy of various popular folk theories that are suggested to YouTubers in order to increase the probability of their videos going viral by utilizing a hypothesis-driven approach.

To this end, we will need to:

- Identify what are the most common folk theories related to the virality of YouTube videos from reputable sources.
- Test these folk theories by evaluating the reliability of the metrics propounded by them using actual YouTube trending video statistics data.

To achieve this objective, we shortlisted and reformulated folk theories from reliable sources. We expected folk theories to be objective and precise, but in many cases had to derive the 'technical/objective' counterpart of the folk theory by looking at multiple similar folk theories, since they were usually vague. For example many sources suggested to keep the video short, but they either didn't specify the video length or they had varying estimates, therefore we had to refer to multiple sources to come up with our estimates. Though our hypothesis testing, we have identified statistical estimates which will now give us a better understanding of the extent to which each of the inferences are true. We initially figured these estimates would allow us to set a threshold to classify folk theories as valid or invalid, but due to the subjective nature of many of our hypotheses, we could not draw a hard classification of valid and invalid for many folk theories.

Outcomes achieved by the project:

- Statistical inference and analysis stating which folk theories are credible.
- Effective communication of the results from folk theory testing using either tabular or graph format.

5. Description of Work Accomplished: Data

The YouTube data we worked on was acquired through Kaggle[10]. The dataset contains metadata for the top 200 most trending videos daily for 7 months. YouTube uses a combination of factors including the measurement of user interactions (shares, comments, likes) to determine the day's top trending videos. The dataset contains data for 10 countries. We've decided to focus on the most trending videos in the United States. The features of the dataset include 16 columns with the following info:

video_id (String)	trending_date (Date)	title (String)	channel_title (String)
publish_time (Date)	category_id (Integer)	views (Long)	likes (Long)
dislikes (Long)	comment_count (Long)	thumbnail_link (String)	comments_disabled (Boolean)
ratings_disabled (Boolean)	video_error_or_removed (Boolean)	description (String)	tags (String)

Since many viral videos trend for many days continuously and may still become popular after a break, we end up with metadata of 6351 unique videos, published over the years 2017 to 2018.

For folk theories, we researched multiple sources through google searches. We shortlisted 12 sources in total [11-24]. We tried to ascertain the credibility of these sources by individually looking into the background of each of the sources. For sources such as Forbes and the Guardian, the reason why their folk theories were highly ranked was obvious since they were reputed journalistic organizations with a history for writing quality articles, but we found highly ranked articles from other authors who we did not recognize readily, for these sources we manually searched and verified their credibility. For example: When we researched Jeff Bullas who is the author of one of our sources, “3 Reasons Why YouTube Videos Go Viral [14]”, we found that he is the owner of jeffbullas.com, and that Forbes calls him a top influencer of Chief Marketing Officers and the world's top social marketing talent. Entrepreneur lists him among 50 online marketing influencers to watch. Inc.com has him on the list of 20 digital marketing experts to follow on Twitter. Another example is Karen Cheng [13], who is the content creator of a viral video herself and in the blog details her approach to making her video go viral.

Once we shortlisted the sources, we had to shortlist the folk theories from these sources. We found that in many cases there was significant overlap, and so for these folk theories, we had to reformulate our folk theories. For example, many authors suggest to have videos to be of shorter duration, but the folk theories would conflict in the exact number or range (in minutes) that they suggest. For example Karen Cheng [13] doesn't give an exact figure, but mentions that a video which is less than 3 minutes would have a better bet of going viral that it's longer counterparts, she ends the folk theory my mentioning that she aimed to get her video down to 1 minute 51 seconds, so her folk theory suggests keeping videos from 2 to 3 minutes. The article published by the Guardian [16] suggested that we needn't keep the videos too short and mentioned that the most viral video of 2015 was 3 minutes long. On analysis for the same folk theory (on time) from all our sources, we decided on the range of 3 to 5 minutes. Therefore we compared the various folk theories and based on the number of times a certain range was repeated selected that as our range for the time of the video.

6. Description of Work Accomplished: Approach

The approach taken to determine the efficacy of a hypothesis was dependent on whether the features necessary for hypothesis testing were present in the dataset and based on the nature of the hypothesis. If the features were present, we performed “Complete Analysis”, else we performed “Subsampled Analysis”.

For the Complete Analysis, we created scripts that would analyze the entire dataset, consisting of viral videos from December 2017 to June 2018. Some videos were viral on multiple days; therefore, we first removed all duplicate video entries from the dataset, and only preserved the latest entry. On this new dataset, we ran a Python scripts to analyze the relationship between the highlighted feature in the folk theory and the number of views the video had got.

For the subsampled approach, we again removed all duplicate entries of the videos, and preserved only the latest entry. Then for each month, beginning from December 2017 till June 2018, we sorted the viral videos of each month in decreasing order of the number of views and picked the top 10 videos of each month. Of the selected 70 videos, three of them were discarded, because the video was no longer available on YouTube. An example of this, is Fergie’s 2018 performance of the national anthem at the NBA all-star game, which was widely panned and was hence taken down by Fergie. We then manually labelled the 67 videos, on 9 different features (Table 1).

Most of the features that we were looking for were easily understandable and objective, like length of the video, and did not require much discussion, but there were others features for which we had to discuss the measurement metric. For example, for whether or not a video starred young individuals. We decided that we would be looking at the main individuals in the video (the stars) and would search for their age. For videos, where there were no human performances, but mere voice overs (common in many science videos), we looked at the age of those who created the video and gave their age. For whether the video ran smoothly in all platform, we tested the video in three different scenarios; on our laptops, on our mobile phones with Wi-Fi and on our mobile phones with a data connection (e.g. 4G). For the feature looking at whether the video was published in a popular channel, we looked at the number of subscribers and number of views the channel had as a whole and counted the channel as being popular if it had more than

100,000 subscribers and had a total number of views that were more than 10 million. For the presence of musical quality, we checked whether music played an important part of the video. This would be the case for music videos but also for movie trailers where the background music is important for the complete experience. However, if there was a video where the music was limited just to the first few and last few seconds, we would not consider it. For whether the video was engaging and entertaining, we tried to think from the perspective average user, whether they would have found the video entertaining. For example, a music video would be engaging and entertaining, but a video on suicide, while important may not be entertaining, and hence was not considered to be so.

We marked 1 if the hypothesis was true for particular video and 0 if it were not. In addition to that we also had an explanation/description column where we filled why we marked it as a 0 or 1. For example, in the video regarding the presence of young individuals, we would mention the age of the stars/video creators in the description column.

Length of the video	Presence of Comedy	Pre-rehearsed video
Presence of Musical Quality	Presence of young individuals	Title in third person
Runs smoothly in all platforms	Engaging and entertaining	Published in popular channels

Table 1: Shortened version of the hypotheses that were looked into in the subsampled analysis (full version is present in the Results section)

7. Description of Work Accomplished: Results

7.1 Complete Analysis

7.1.1 Viral videos are published on Monday or Tuesday

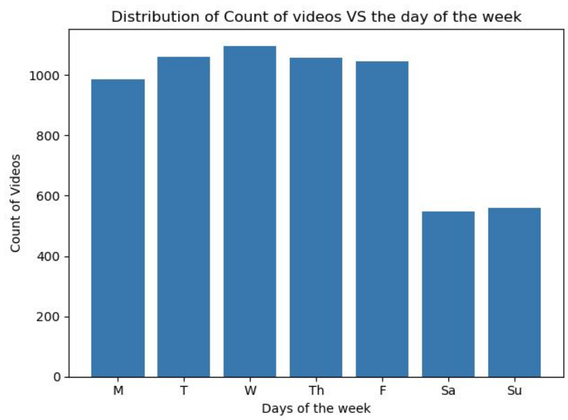


Figure 1 : Distribution of count of videos vs the day of the week it was published

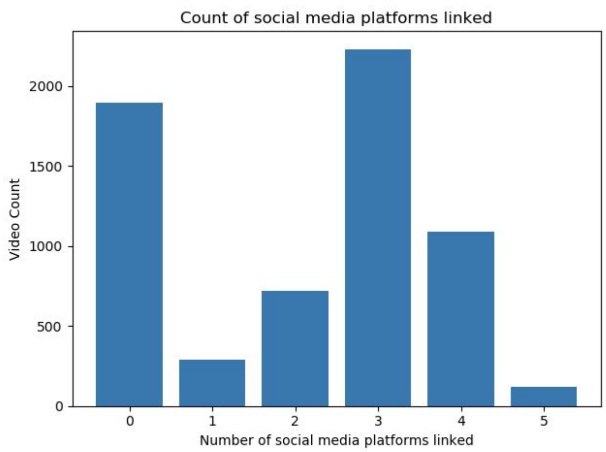
The reasoning behind the hypothesis was that if the video was published at the beginning of the week, people would watch it during their work. This hypothesis assumed that the per-day view count of the videos will reduce during weekends.

Our analysis found that this hypothesis was mostly wrong, in that the viral videos are published throughout the week almost uniformly, except for the weekends. On weekends there is a marked drop in the number of published videos, almost by half. Therefore seems to be some merit in the assumption that videos may not go viral over the weekends(Figure 1).

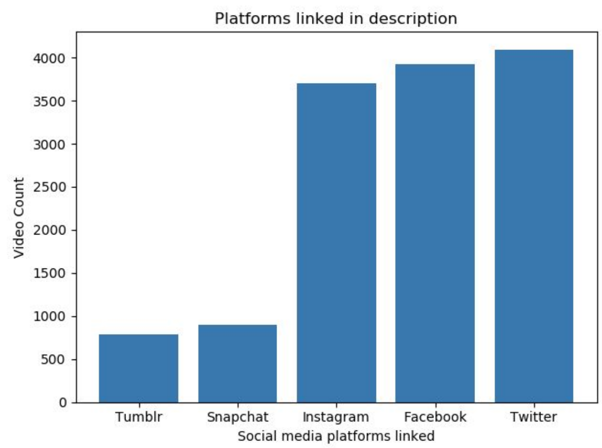
7.1.2 Viral videos are shared/linked to other Social Media Platforms

The folk theory corresponding to this analysis was that for a video to go viral the content creator has to share the video on other social platforms and also be active in marketing the video on those platforms. Therefore the hypothesis testing we've done is to search for the attempt to connect via other social media accounts. Particularly we search if their Twitter, Facebook, Tumblr, Snapchat and Instagram accounts are linked in the description and the count of the different social media accounts being used.

As shown in Figure 2a, our analysis found that there is a larger number of videos whose content creators link to at least one other social media account, this gives them the advantage to market their video on those platforms as well and be allowed to keep in touch with audiences for a greater view count for their upcoming videos. As for the platform that is most widely linked in viral videos, it's Twitter, followed by Facebook and Instagram, as shown in Figure 2b.



a) The number of linked social media accounts



b) The number of videos in which a particular social media platform is used

Figure 2: Relation between the virality of videos and the tendency of YouTubers to link other social media accounts

7.1.3 Viral videos go viral within a day of being published

We found that a significant percentage (though not a majority) of the videos in our dataset did go viral within a day of being published. 44.9% of videos became viral within one day of it being published and 73.5% of the videos become viral within two days of the video becoming published(Figure 3).

7.1.4 There exists high correlation between the number of comments and the number of views on the viral video

As seen in Figure 4, we found that comments and view count have a positive correlation with a Pearson correlation value of 0.59 ($p<0.05$) for viral videos.

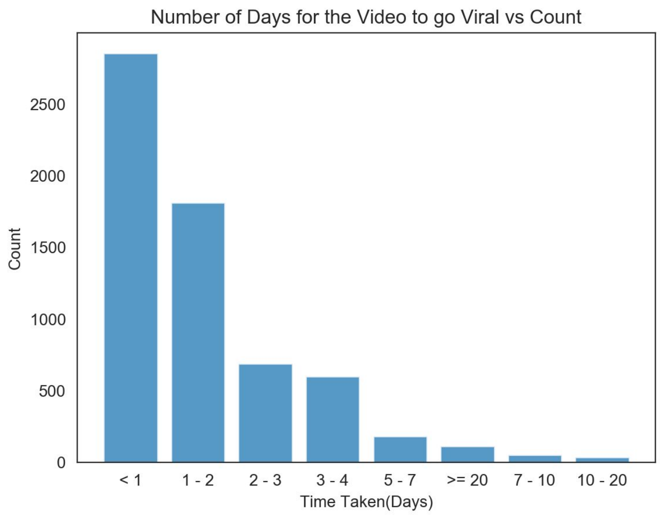


Figure 3. Time taken for the video to go viral

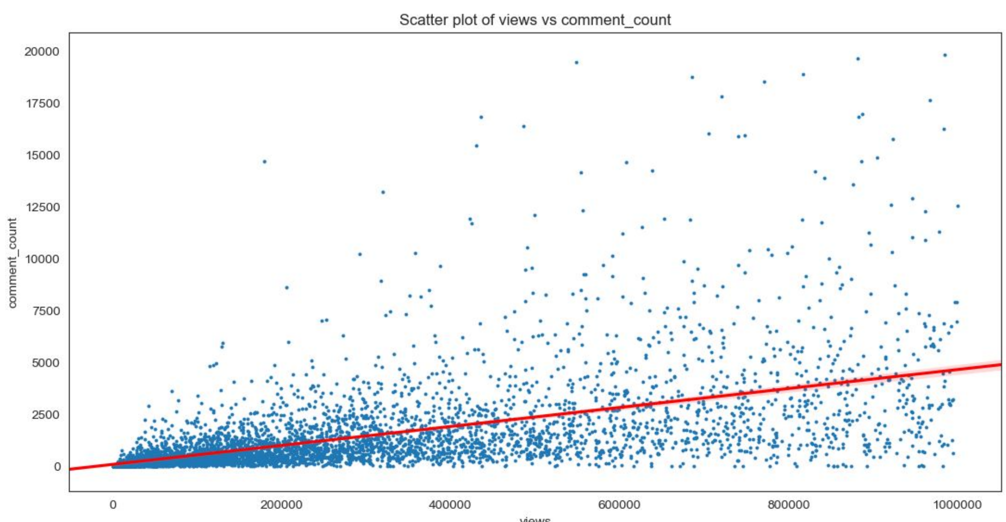


Figure 4: Correlation between comments and views for viral videos

7.1.5 Viral videos have short video titles(≤ 4 words)

We found that the majority of the videos in our dataset had an average of 7-8 words (Figure 5) in their title.

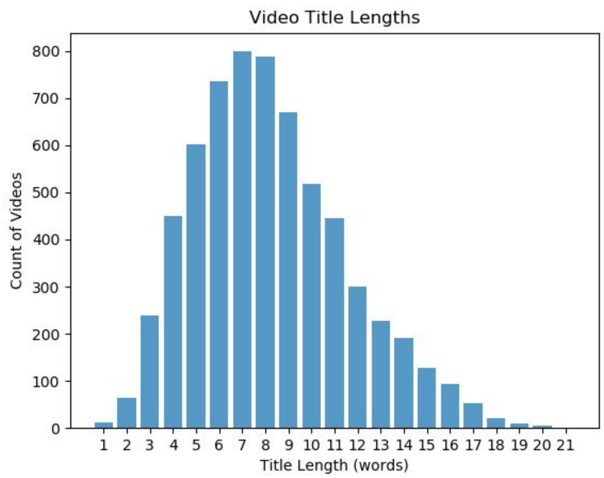


Figure 5: Title length of viral videos

7.2 Subsampled Analysis

Hypothesis	Number of subsampled videos that agreed to the hypothesis	Percentage of subsampled videos that agreed with hypothesis
Viral videos are short(<=300 seconds)	51/67	76%
Comedy genre videos are more likely to go viral	8/67	12%
Viral videos are rehearsed and not candid	61/67	91%
Viral videos have at least few musical elements embedded in the video(can be background music or the entire video can be a music video)	61/67	91%
Viral videos have the majority of individuals in the video under the age of 35 years	54/67	82%
Viral videos titles are written in third person	64/67	96%
Viral videos run smoothly on all platforms and devices(especially mobile phones)	67/67	100%
Viral videos are generally engaging or entertaining	63/67	94%
Viral videos are generally published by popular YouTube channels	66/67	98%

We have discussed the efficacy of the different hypotheses during the Subsample analysis and highlighted counter examples from the same dataset.

7.2.1 Viral videos are short (< 300 seconds)

76% of the videos in our subsampled analysis agreed with the hypothesis. People are getting impatient and not watching long videos. A counter-example is the video, “YouTube Rewind : The Shape of 2017”, which was more than 7 minutes long and yet went viral since it was entertaining and at the same time had all the top YouTubers from the year in a single video.

7.2.2 Comedy genre videos are more likely to go viral

From our analysis we were able to conclude that videos belonging to the Family genre are most popular viral videos and only 12% of our subsampled analysis were actually funny(Figure 6).

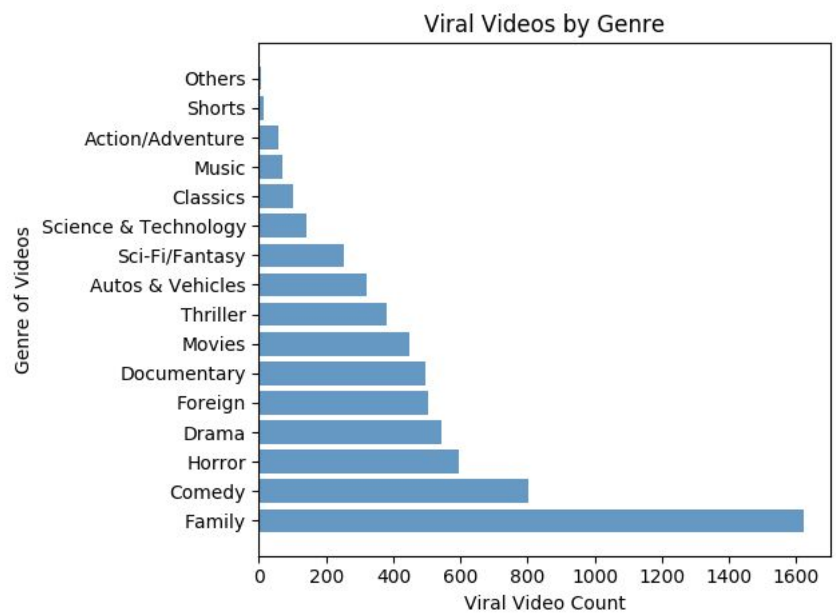


Figure 6: Genres of viral videos

7.2.3 Viral videos are rehearsed and not candid

It is clear from our analysis that more than 90% of the sampled video were properly rehearsed and were heavily edited. However, this is very different from TikTok, another popular video sharing platform where users upload 3 to 15 second candid videos. This shows that it is possible for candid videos to go viral, but that this usually does not occur on YouTube.

7.2.4 Viral videos have at least a few musical elements embedded in the video (can be background music or the entire video can be a music video)

The bulk of the videos (91% of the subsample) have musical elements in the video. We noticed that scientific videos, news and speeches generally do not incorporate any musical elements in them.

7.2.5 Viral videos have the majority of individuals in the video under the age of 35 years

54 out of the 67 videos we analysed had most of the individuals under the age of 35 years. However, we did come across a video by singer-actress Jennifer Lopez titled “El Anillo” which comprised of actors above the age of 35 years.

7.2.6 Viral videos titles are written in third person

We never expected such a folk theory while surveying the available resources but surprisingly 96% of our subsample actually agree to the hypothesis. A counter to this theory is the viral video “To Our Daughter” published by Kylie Jenner(famous online influencer) that talks about her pregnancy.

7.2.7 Viral videos run smoothly on all platforms and devices(especially mobile phones)

All videos in our subsample agreed to this hypothesis. We believe that this hypothesis would have been more relevant during the early years of YouTube when some of the published videos could not render properly on both large screens and small screens even when there was a stable internet connection.

7.2.8 Viral videos are generally engaging or entertaining

The general expectation is that only engaging and entertaining videos go viral. Contrary to popular belief, videos with a strong emotional message though bland go viral regularly - one such example is the video uploaded by the YouTuber Logan Paul spreading sucide awareness(Suicide: Be Here Tomorrow).

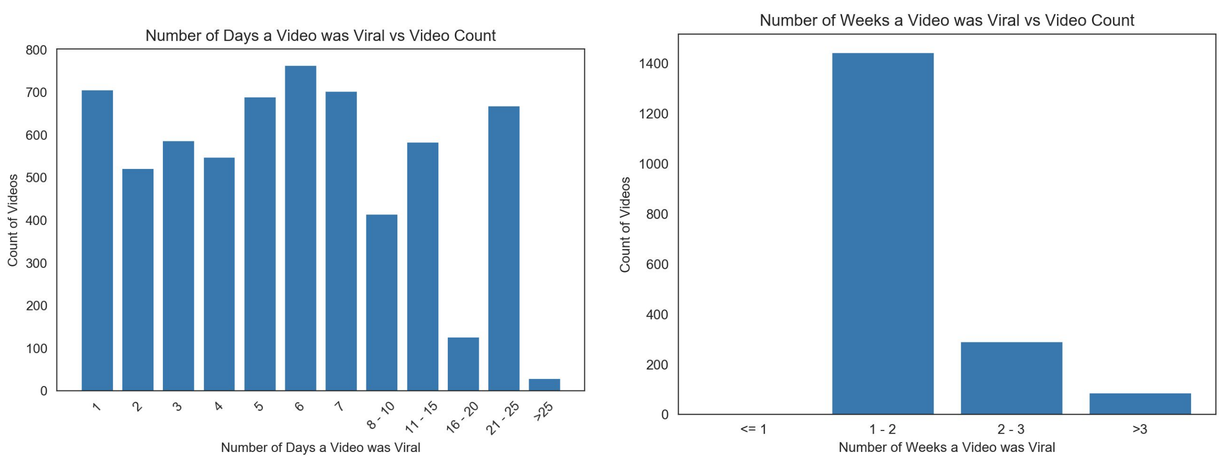
7.2.9 Viral videos are generally published by popular YouTube channels

We were able to find only one video in our subsample which was published by an unpopular YouTube channel. “5 Senses with Dr. Oz” published by Turkish Airlines which went viral with 87 million views in a span of a few days. Out of 150+ videos uploaded by this channel, this is the only video that went viral and aligns with a number of previously mentioned hypothesis. The video is short, has a background score, entertaining, engaging and not to forget has the presence of Dr. Oz, who is a renowned celebrity, in the video.

7.3 Supplementary Analysis

7.3.1 Longevity of Viral Videos

As a general analysis, we measured how long the videos in our dataset remained viral. We found that if we measured viral longevity of the videos by the number of days, it produced varied results as shown in Figure 7a. However, if we measured viral longevity by the number of weeks, we found that most videos remained viral for 1-2 weeks as shown in Figure 7b.



a) Viral longevity measured by number of days

b) Viral longevity measured by number of weeks

Figure 7: Measure of how long videos remained viral

7.3.2 YouTuber Collaborations

This folk theory believed that if two or more YouTubers collaborated on a single video it would bring together subscribers from each of their seperate channels and the chances of their video going viral would be higher. We are testing this hypothesis by analyzing the title of the video, and looking for the presence of keywords like ‘feat’ and ‘featuring’ which are mostly mentioned in the title of collaboration videos. We found that there were a negligible amount of collaboration present in the videos in our dataset.

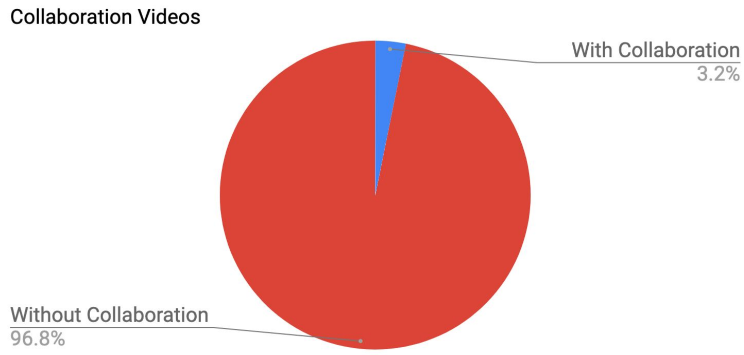


Figure 8: Percentage of viral videos that featured collaborations by two or more YouTubers

8. Discussion of Outcomes, Implications, and Conclusion

8.1 Discussion and Conclusion

Through our study we observed that many of the popular folk theories that are present on the internet do indeed turn out to be true. However it should be noted that some of the theories do not play out in reality or are only partially true. This points to the important fact that these theories should not be blindly trusted. It should be noted that the folk theories we tested were chosen from reputable sources (eg: Guardian) or experts. It is indeed possible that if theories from less reputable sources were instead tested, the number of flawed or misleading hypothesis would be substantially more.

Our study also shows that manual annotation of even a small sample of the videos (when the data is already available), can lead to interesting insights and provide a more holistic picture of the phenomenon we are studying.

Some of the findings were in line with our expectations. For example, we expected viral videos to be entertaining and engaging as we viewed the platform as primarily an outlet for users to have fun and relax and largely not for pursuing serious conversations and discussions. Additionally, we thought that the viral videos would indeed mostly feature young individuals since we thought that they would reflect the demography that mostly uses the platform (namely those who are millennials or younger).

However, there were some findings that surprised us. For example, we did not expect viral videos to favor using titles in third person as opposed to first and second person. We had assumed that the titles would be equally likely to be used in our styles as there did not seem to direct relation between the style of writing and virality of a video. Additionally, it was interesting to see such as stark difference between the number of viral videos being published on weekdays vs weekends. If anything we expected there to be more viral videos published on weekends because viewers would potentially be more free to watch videos.

8.2 Relevancy of the Project to the Course

The specific success metrics YouTube uses in its curation algorithm to determine if a video is trending is unknown. However, the findings in this project gives us an idea about what types of video content typically trend on YouTube. As discussed in this course, if curation algorithms are too personalized for users, they run the risk of creating "filter bubbles" and promoting polarization on their corresponding online platform. Even though the display of viral videos on YouTube are not personalized, it still has the disadvantage of occasionally popularizing the same types of content even if there are other videos that users may potentially like more.

Our team was introduced to the concept of folk theories through this project, and our entire project is centred around folk theories related to the virality of YouTube videos. Folk theories is a concept of social computing systems that were not taught in class, and we learnt it through the project.

8.3 Limitations and Future Research Directions

The project evaluated the data for only one region or one country. Similar research can be performed in other countries, and analysis of variance(ANOVA) can be carried to determine similarities in the distributions.

The analysis was restricted to only about seven months of data. One can look at analysing the data over a longer time frame using autoregressive integrated moving average(ARIMA) time series analysis. We could look at how the validity of specific folk theories has changed over time.

While searching for folk theories online, many of the suggestions were catered towards marketing and political advertisements. We ignored the folk theories related to marketing since it was out of the purview of the project. Also, there were no upcoming elections in the USA, and hence our dataset had few videos related to politics. Thus, we did not evaluate politically-themed folk theories.

The analysis in the project concentrated more on the metadata we have on the viral videos and not the video content itself. Analysing the video content and user interactions in the comments section can provide more meaningful insights on viral videos and their characteristics.

The Subsampled Analysis was carried out by our team members. We may have indirectly projected our bias into the results while evaluating individual videos. This problem can be better handled by using services like Amazon Mechanical Turk and asking third parties to assess the videos to minimise the bias.

The entire study was carried out on viral videos available on YouTube. For a more comprehensive analysis, one can perform a similar analysis on videos published on platforms like TikTok, Vimeo, Instagram, Facebook, DouYin, etc. One can also check the validity and applicability of each folk theory on these different platforms.

There is no absolute definition of virality, and there are different curation algorithms that can be utilised to determine the viral videos. Our dataset used YouTube Trending Videos API to obtain the data. For us, the YouTube curation algorithm is a black box, and we are not able to objectively explain why a particular video was categorised as viral. A different curation algorithm may not consider the same video as viral. Further research can be carried out in understanding the different curation algorithms and their impact on video virality.

List of Team Members

Name	**Email**	**GT ID**
Arvind Akpuram Srinivasan	arvind_s@gatech.edu	903528961
Karthik Nama Anil	kanil3@gatech.edu	903471605
Miasia J Jones	mjones386@gatech.edu	903164042
Prithvi Alva Suresh	al.prithvi@gatech.edu	903541747

References

[1] Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., & Moon, S. (2007). I tube, you tube, everybody tubes. *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement - IMC 07*. doi:10.1145/1298306.1298309

[2] Guadagno RE, Rempala DM, Murphy S, Okdie BM: What makes a video go viral? An analysis of emotional contagion and internet memes. *Comput Hum Behav* 2013, 29(6):2312–2319. 10.1016/j.chb.2013.04.016

[3] Tyler Wes (2011). Going Viral: Factors That Lead Videos to Become Internet Phenomena.*The Elon Proceedings of Journal of Undergraduate Research in Communications Vol. 2, No. 1*

[4] Lewis Howes (2012). How To Go Viral On YouTube: The Untold Truth Behind Getting Views. https://www.forbes.com/sites/lewishowes/2012/08/09/how-to-go-viral-on-youtube-the-untold-truth-behind-getting-views

[5] Mackenie Graham (2015). 9 tips to make a video go viral on YouTube. Retrieved from: https://thenextweb.com/insider/2015/11/20/9-things-to-help-your-youtube-video-go-viral/

[6] Adam Toren (2015). 5 Ways to Boost Your Chances of Going Viral on YouTube. https://www.entrepreneur.com/article/247661

[7] Feroz Khan, G. and Vong, S. (2014), "Virality over YouTube: an empirical analysis", Internet Research, Vol. 24 No. 5, pp. 629-647.

[8] OpenRefine. Retrieved from: http://openrefine.org

[9] Wrangler. Retrieved from: http://vis.stanford.edu/wrangler/

[10] Trending Youtube Video Statistics. Retrieved from: https://www.kaggle.com/datasnaek/youtube-new

[11] 10 ways to make your video go viral. Retrieved from: https://medium.com/this-happened-to-me/10-ways-to-make-your-video-go-viral-d19d9b9465de

[12] Secrets of YouTube – what makes a video go viral . Retrieved from: https://www.theguardian.com/voluntary-sector-network/2015/nov/05/youtube-what-makes-video-go-viral-charities

[13] Medium. (2019). 10 ways to make your video go viral. [online] Available at: https://medium.com/this-happened-to-me/10-ways-to-make-your-video-go-viral-d19d9b9465de [Accessed 9 Dec. 2019].

[14] Bullas, J. (2019). *3 Reasons Why YouTube Videos go Viral*. [online] Jeffbullas's Blog. Available at: https://www.jeffbullas.com/3-reasons-why-youtube-videos-go-viral/ [Accessed 9 Dec. 2019].

[15] Biteable. (2019). *How To Make A Video Go Viral: The Magic Formula | Biteable*. [online] Available at: https://biteable.com/blog/tips/how-to-make-a-video-go-viral/ [Accessed 9 Dec. 2019].

[16] Upscope Blog!. (2019). *Secrets of YouTube: What makes a video go viral*. [online] Available at: https://blog.upscope.io/secrets-of-youtube-what-makes-a-video-go-viral/ [Accessed 9 Dec. 2019].

[17] The Guardian. (2019). *Secrets of YouTube – what makes a video go viral*. [online] Available at: https://www.theguardian.com/voluntary-sector-network/2015/nov/05/youtube-what-makes-video-go-viral-charities [Accessed 9 Dec. 2019].

[18] Forbes.com. (2019). *How To Go Viral On YouTube: The Untold Truth Behind Getting Views*. Available at: https://www.forbes.com/sites/lewishowes/2012/08/09/how-to-go-viral-on-youtube-the-untold-truth-behind-getting-views/#138e47c16b97 [Accessed 9 Dec. 2019]

[19] Inc.com. (2019). *Want Your Video to go Viral? The Rules Have all Changed*. [online] Available at: https://www.inc.com/christina-desmarais/5-steps-to-a-viral-video-according-to-a-guy-behind-youtubes-number-one-ad-of-decade.html [Accessed 9 Dec. 2019].

[20] Harvard Business Review. (2019). *Why Some Videos Go Viral*. [online] Available at: https://hbr.org/2015/09/why-some-videos-go-viral [Accessed 9 Dec. 2019].

[21] Forbes.com. (2019). *6 Qualities To Make Your Videos Go Viral*. [online] Available at: https://www.forbes.com/sites/ilyapozin/2014/08/07/6-qualities-to-make-your-videos-go-viral/#18a80271154e [Accessed 9 Dec. 2019].

[22] Wyzowl. (2019). *How to Make a Video That is Guaranteed to Go Viral*. [online] Available at: https://www.wyzowl.com/how-to-create-a-video-that-is-guaranteed-to-go-viral/ [Accessed 9 Dec. 2019].

[23] Toren, A. (2019). *5 Ways to Boost Your Chances of Going Viral on YouTube*. [online] Entrepreneur. Available at: https://www.entrepreneur.com/article/247661 [Accessed 9 Dec. 2019].

[24] Sibley, A. (2019). *Common Qualities of Insanely Successful Viral Videos*. [online] Blog.hubspot.com. Available at: https://blog.hubspot.com/blog/tabid/6307/bid/33357/common-qualities-of-insanely-successful-viral-videos.aspx [Accessed 9 Dec. 2019].

[25] Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First I "Like" It, then I Hide It: Folk Theories of Social Feeds. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2371–2382.