

Research on Prediction of Different Categories of Video based on YOUTUBE Using Text Mining and Sentiment Analysis

Lin Gan *

Department of Computing, the Hong Kong Polytechnic University, Hong Kong, China

* missmissganlin@gmail.com

Abstract. With the development of online commentary research, scholars have tried to tap into the deeper value of online commentary from the analysis of sentiment analysis, quality evaluation, false comment recognition to the usefulness of comments. Previous studies have focused on online product reviews while news reviews. Social media research has been relatively rare. social media and news commentary contain readers' opinions and evaluations on current events, and reflect the trend of public opinion. The purpose of this paper is to investigate and analyze the intrinsic link between social media content of different type and the number of commentaries, and sentiment analysis.

Keywords: Social Media; Video; Text Mining; Sentiment Analysis.

1. Introduction

There are many social networks in use today, like Facebook, Twitter, YouTube, and many others. Social media has been broadly used in the recommender system, sentiment analysis, text mining, business approach and marketing. A key aspect of social networks is richer in information that brings the challenges and opportunities in the data mining and knowledge discovery. There makes it appear that many researchers have started to study online reviews and have already made good progress. For example, according to the usefulness of online product reviews, the reviews can be sorted and then followers Can get comments with more valuable information, which makes easier to complete next step decisions. Further, most social media researches focused on excavating the emotional factors contained in commentary, evaluating the quality of different content commentary.

2. Literature Review

In 1999, Leets et al. (1995) explored the reasons why people sent letters to celebrities. Nowadays, a lot of people and media join the social media platform and frequently interact with their followers. In this study, we aim to investigate the underlying reasons of people leaving their comments on different channels of videos. According to comments based on Leets et al. (1995)'s categorization of why people Interact with celebrities. we got an idea that collected different YouTube sample comments. This helps us to explore this kind of comments at a larger scale. The collected data will be summarized and discussed. The following hypotheses are thus drawn accordingly. Are the number of comments on social media directly related to video contents?

H₁, Whether the categories of videos affect the number of comments on social media?

H₂, Whether the viewer's comments directly related to the videos content? (positive/negative).

(This paper extracts text features, sentiment features, data analysis from different type videos reviews according to relevant research, formulates an evaluation system for video review and video categories popularity prediction task.).

3. Methods

We collected six topics videos on the YouTube hot search list from February 31. January 2019 to.

The first step is choosing the top 6 popular categories of videos on YouTube, like Haul video, Online teaching, VLOGs, Game video, Comedy /Skit video and Unboxing Videos.

The second Step is collecting each category 8-10 samples in same categories of channel. Moreover, each channel choose 6 equally number of viewers for comparison.

Table 1. Online teaching INCLUDING Beauty (makeup) & Cooking

Channel name	More Than 15M views	5M viwes	More than 1M views	900Kviews	300K views	100K views
Bethany Mota	v=koMbIaJ8Tmo Morning Routine: Fall Edition!!	v=EoakbXJmcto Fears, TV Shows, Boyfriend? TMI TAG!	v=0K5XdYe4EGA name: Instagram Followers Control My Life for a Day Bethany Mota 1693	v=PbH8oC_OyVM Name: Summer Try-On Clothing Haul! Forever 21, Urban Outfitters, & more!	v=bREL_G2Y46g name: \$100 Thrifted VS \$100 at Urban Outfitters! Clothing Haul 928	v=I1qczcjh9XM name: Spring Jewelry - My Must Have Pieces! 628
Michelle Phan	v=qg-2rDnWCJA Seductive Vampire	v=zupy7R4uR2g Seductive Smokey Eyes Tutorial	v=MXm7obi1yHM Final Dynasty : Aesthetic Inspiration 1713	v=swvvrpNIPTc Trend Report: Glitter Freckles, Geometric Undercuts + more 1214	v=cYSYnWwy3R8 Pillow Talk 8: Cyber-Bullying and Acne Shaming 731	N/A
Rachel Levin	N/A	v=eGhsWJgbuzg Doing My Makeup In 5000 Pounds of Jello With James Charles!	v=l8W_RpxDa3I Swapping Outfits With Tana Mongeau! 1964	v=XNxoeO_yU14 Christmas Carpool Ride! 2108	v=MhW8pL2jx10 Handcuffing My Boyfriend To Josh Peck For A Day! 1644	v=-_sw3o5Szcs 5 Cute Valentines Nail Designs
Nikkie De Jager	v=SOoauHBpcbU Side by Side + Jessie J	v=fUgxywvVHc How to Contour & Highlight	v=CGJkvTjoV5Y TRANSFORMING MY MOM INTO ME! 2082	v=QMQR0Httcrc Glittery Double Cut Crease & Purple Lips Holiday Makeup 1815	v=KmZZGfntdQw SNAPCHAT Q&A: Huda Beauty, Getting Married & MORE! 1166	v=IpoUjFgc-2g Edgy Flapper Girl Makeup Tutorial · RuPaul's Drag Race Series 272
food wishes	N/A	v=uxEhH6MPH28 Buttermilk Fried Chicken	v=UcqzDbB-9tI Lemon Pepper Potatoes - Food Wishes 1117	v=71ZJRshrEJM Butcher's Steak (aka Hanger Steak) - How to Trim and Cook Butcher's Steak 986	v=3IZFd014vQM German Potato Dumplings (Kartoffelkloesse) - Food Wishes 1166	v=l6Q2QMbEgZk Chef John's White Bread - Food Wishes 726

Munchie	N/A	N/A	v=eUPYFt52F BE How To Make Jamaican Beef Patties 1296	v=0fxL8v2dMh o How to Sharpen a Knife with a Japanese Master Sharpener 350	v=iGtqF3N6w CI The Capital of Cecina Tacos - Ultimate Taco Tour of Mexico 245	v=Q7OqC0qtJ Yc How War Changed Korean Food Forever with Maangchi & Japanese Breakfast 425
Binging with Babish	N/A	v=UvlkojfQDx A Binging with Babish 1 Million Subscriber Special: Taco Town & Behind the Scenes	v=_SOBosXrD Rc Pulled Pork Basics with Babish 1776	v=iF2yGGfGv1 8 Sous Vide Basics with Babish Live 222	v=nKKznGaN7 PI Tacos Basics with Babish Live 197	v=3S94NIJ8o QA Quesadillas Basics with Babish Live 156

According to each channel video ID (e.g., V=23916491239) to mining the online commentary. Here are some example and outcome below:

Positive:37.21075672295184

Negative:62.78924327704816

Fig 1. Example 1

I AM NEVER COLLABORATING WITH YOU AGAIN THIS WAS THE WORST
 Joe "you can put your meat in me" Tyler "ooohhh yes''XD
 "What are the lyrics to Troye's song?" "COCAINE."
 Joe: "You can put your meat in me.."
 Tyler: "Omg yes 😊"
 Hahahahhaa ffs tyler
 "YOU CAN PUT YOUR MEAT IN ME, TYLER" - Joe Sugg
 IS ANYONE ELSE IN LOVE WITH TYLERS LAUGH???(😂😂)
 Tyler no one on their right mind would put a baby in a fridge 😊
 I'm really dirty minded so this was just too much
 I lost it when tyler said a flashlight 😊
 tyler shouting COCK had me laughing SO hard im crying xD
 Tyler Oakley is my spirit animal seriously tho 🐶🐶
 I love you and Tyler!!!
 Love Tyler's laugh soooo much😂😂❤️
 Love Tyler's laugh soooo much😂😂❤️
 Tyler has a magical laugh
 Did anyone notice tyler left a red mark on joes chest by slapping him in the
 beginning
 I ship them hardcore
 "you can put meat in me!"- Joe sugg 2014
 the noise Tyler makes at 4:25 is so suggestive 😊😂 and so are the noises he
 makes that follow it 😊
 PLEASE MAKE MORE OF THIS !!!!! 😂😂😂
 Joe's chest was so red after that slap at 0:41
 WHAT IF DAN AND PHIL DID THIS

Fig 2. Outcome 1

After I collected six categories of data, we used measured variables and the correlations between and measured variables and influence of viewers activity variables was conducted first. To exam the relationship between the measured independent variables in predicting emotional response, structural equation modelling was conducted.

4. Result

The research has three main findings, firstly, the hot topic keywords of public attention in different entertainment videos are slightly different, in addition, the public's positive emotions for Haul videos and unboxing videos with positive emotions increasing as whole. Public's comments are more positive and excited. Second, game video is the most popular category of video. Finally, the number of video views fluctuated greatly with the upload time, the number of viewers reached the peak in the week after the latest video uploaded in the first stage, in the second phase, the number of people dropped rapidly form the peak during a period, and third phase entered a slow decline period, until the end, the number of visitors is fixed at a constant value.

5. Conclusion

In the research, we analyzed different categories of video and proposed new method of social media comments to analyze the viewers' needs. This method uses gender as an example to analyze viewer preference. Through empirical analysis of different categories of video, it is found that female viewers prefer videos such as Haul videos, unboxing videos and beauty teaching (online teaching), while male viewers pay more attention to games and technology-related unboxing videos.

Comments affected by video content, communication technology and media environment. Among the above six categories of video, we conclude that there are a lot of comments on game videos, but the negative sentiment is very high. Because viewers cannot take part in the real game, thus, their experience of immersive gameplay cannot be obtained.

References

- [1] Ghani, Norjihhan Abdul, Suraya Hamid, Ibrahim Abaker Targio Hashem, and Ejaz Ahmed, "Social media big data analytics: A survey", *Computers in Human Behavior*, 2018.
- [2] Sapountzi, Androniki, and Kostas E. Psannis, "Social networking data analysis tools & challenges", *Future Generation Computer Systems*, Vol.86, pp.893-913, 2018.
- [3] Tanev Hristo, Bruno Pouliquen, Vanni Zavarella & Ralf Steinberger (forthcoming). Automatic Expansion of a Social Network Using Sentiment Analysis. In: *Annals of Information Systems, Special Issue on Data Mining for Social Network Data*. Uspensky, B. (1973). *A Poetics of Composition*. University of California Press, Berkeley, California.
- [4] Mazer, J., Murphy, R. and Simonds, C. 2007. I'll see you on 'Facebook': The effects of computer-mediated teacher self-disclosure on student motivation, affective learning, and classroom climate. *Communication Education*, 56(1): 1-17.
- [5] Selwyn, N. Screw blackboard... do it on Facebook!: An investigation of students' educational use of Facebook. Paper presented to the Pole 1.0 - Facebook social research symposium. November15. University of London. <http://www.scribd.com/doc/513958/Facebook-seminar-paper-Selwyn> (accessed September, 2008).
- [6] Tufekci, Z. 2008. Grooming, gossip Facebook and Myspace. *Information, Communication & Society*, 11 (4): 544-64.
- [7] Wellens, J., Hooley, T., Meek, J. and Madge, C. Forthcoming. "Facebook". In *social integration and differentiation*.
- [8] Lange, P. 2007. Publicly private and privately public: Social networking on YouTube. *Journal of Computer-Mediated Communication*, 13(1) <http://jcmc.indiana.edu/vol13/issue1/lange.html> (accessed December, 2008).