Engagement and Popularity Dynamics of Youtube Videos and Sensitivity to Meta-Data

Mr.T.N.V.S Praveen[1],Sushma Kolli[2],Yamuna Bezawada[3], Renuka Lingam[4],Ramsai Reddy Vintha[5]
[1]Asst.Prof, [2,3,4,5]BTech Students
Department of Computer Science, Lakireddy Balireddy College of Engineering, Mylavaram

Abstract - YouTube is one of the largest leading preferred destinations for watching videos in online. YouTube allows channel creators to monetize their popular videos. The popularity of the videos depends on the meta-level features and social dynamics, i.e. the interaction of the channel creators with YouTube users. The key meta-level features are number of subscribers, first day view count, numbers of keywords used, category of video, title length, Google hits. Optimization of meta-level features after a video is posted increases the popularity of videos. In social dynamics, we discover the causal relationship between views to a channel and subscribers count.

I. INTRODUCTION

The YouTube with 1 billion users who collectively watch millions of hours of YouTube videos and generates billions of views everyday and users upload over 300 hours of video content every minute.

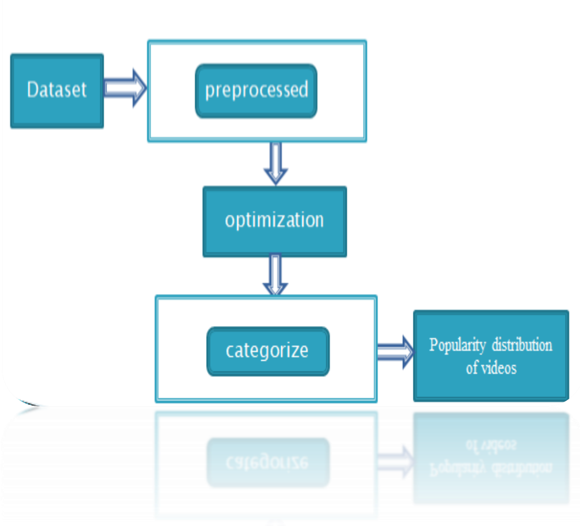
The key metric measure for popularity of videos is view count. A key statement is, How do meta level features of a posted video drive the user engagement in the YouTube social network. Our main aim is to examine how the individual video features contribute to the popularity of a video. The study of popularity of YouTube videos based on meta-level features is the key challenging problem given the diversity of users and channel providers.

The first five key meta-level-features that affect the popularity of a video are: first day view count, number of subscribers, contrast of the video thumbnail, keywords count and Google hits. Optimization of meta-level features after a video has been posted increases the popularity of the video and by optimizing the title increases the traffic due to YouTube search, optimizing the keywords increases the traffic from promoted and related videos and optimizing the thumbnail increases the traffic from related works.

II. PROPOSED SYSTEM

We conducted a data-driven study of YouTube based on a large dataset. By using several machine learning methods, we investigated the sensitivity of the videos meta-level features on the videos view count. The key meta-level features include: first day view count, number of subscribers, Google hits, Keywords count, video category, contrast of video thumbnail, title length and number of upper-case letters in the title respectively. Optimizing the meta-data after the video is posted improves the popularity of the video. The social dynamics i.e, the interaction of the channel also affects the popularity of the channel.

Architecture diagram:



III. METHODOLOGY

We apply machine learning methods i.e, Extreme learning machine to study how video view count impacted by meta-level features of YouTube videos.

We analyse changing of meta-level features, after a video is posted, impacts the user engagement of the video. We study how optimizing the title, thumbnail or keywords affect the view count of YouTube videos. Let ti be the time at which the meta-level optimization was performed on video i and let si, denote the corresponding sensitivity. We characterize the sensitivity to meta-level optimization as follows:

$$s_i = \frac{\left(\sum_{t=\hat{\tau}_i}^{\hat{\tau}_i+6} v_i(t)\right)/7}{\left(\sum_{t=\hat{\tau}_i-6}^{\hat{\tau}_i} v_i(t)\right)/7}$$

The numerator is the mean value of the view count 7 days after optimization. Similarly, the denominator is the mean value of the view count 7 days before optimization.

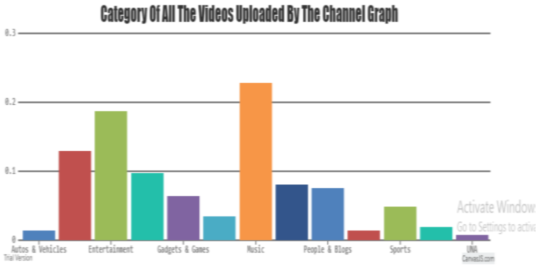
IV. RESULTS AND DISCUSSIONS

The results are shown in the form of a statistical form and graphs like a bar graph, pie chart, line graph, scatter plot, histogram etc. The graph is generated using js chart of predefined libraries, which we can install and execute our graph.

Results for Categorizing:

Category Of All The Videos Uploaded By The Channel

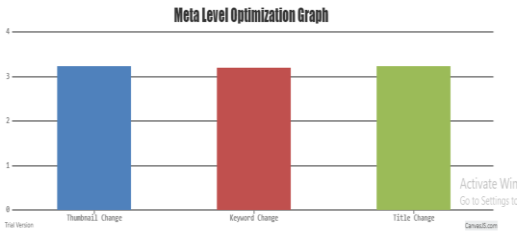
Autos & Vehicles	Comedy	Entertainment	Film & Animation	Gadgets & Games	Howto & DIY	Music	News & Politics	People & Blogs	Pets & Animals	Sports	Travel & Places	UNA
0.014032301	0.1296664	0.18685465	0.09700159	0.064667724	0.0340879	0.22805798	0.080387875	0.07558909	0.013602065	0.04864972	0.019228224	0.008174477



Results for Optimizing:

Meta Level Optimization

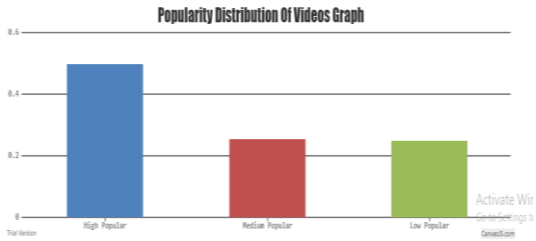
Thumbnail Change	Keyword Change	Title Change
3.225736177884615 4	3.196278160500236	3.222885404056045 3



Results for Popularity of distribution:

Popularity Distribution Of Videos

Highly Popular	Popular	Unpopular
0.4979671	0.25296742	0.24906552



V. CONCLUSION

We conducted a data-driven study of YouTube data using dataset. First, by using several methods(machine learning), we investigated on the view counts of videos for the sensitivity of the videos meta-level feature.It was found that the key meta-level features include: first day view count , Google hits, number of subscribers, contrast of the video thumbnail, keywords count, video category, title length, and number of upper-case letters in the title respectively. We are optimizing the meta-data after the video is posted improves the popularity of the video. The interaction with the YouTube users also affects the popularity of the channel. The final results of the paper was to categorize , to optimize and to know the popularity of YouTube videos.

VI. REFERENCES

[1]. G. G€ursun, M. Crovella, and I. Matta, “Describing and forecasting video access patterns,” in Proc. IEEE INFOCOM, 2011, pp. 16–20.

[2]. H. Pinto, J. Almeida, and M. Gonc¸alves, “Using early view patterns to predict the popularity of YouTube videos,” in Proc. 6th ACM Int. Conf. Web Search Data Mining, 2013, pp. 365–374.

[3]. C. Richier, E. Altman, R. Elazouzi, T. Jimenez, G. Linares, and Y. Portilla, “Bio-inspired models for characterizing YouTube viewcout,” in Proc. IEEE/ACM Int. Conf. Advances Social Netw. Anal. Mining, 2014, pp. 297–305.

[4]. C. Richier, R. Elazouzi, T. Jimenez, E. Altman, and G. Linares, “Forecasting online contents’ popularity,” arXiv:1506.00178, 2015.

[5]. A. Zhang, “Judging YouTube by its covers,” Dept. Comput. Sci. Eng., Univ. California, San Diego, 2015. [Online]. Available: http://cseweb.ucsd.edu/ jmcauley/cse255/reports/wi15/ Angel%20Zhang.pdf

[6]. T. Yamasaki, S. Sano, and K. Aizawa, “Social popularity score: Predicting numbers of views, comments, and favorites of social photos using only annotations,” in Proc. 1st Int. Workshop Internet- Scale Multimedia Manage., 2014, pp. 3–8.

[7]. T. Yamasaki, J. Hu, K. Aizawa, and T. Mei, “Power of tags: Predicting popularity of social media in geo-spatial and temporal contexts,” in Advances in Multimedia Information Processing. Berlin, Germany: Springer, 2015, pp. 149–158.

[8]. T. Trzcinski and P. Rokita, “Predicting popularity of online videos using support vector regression,” arXiv:1510.06223, 2015.

[9]. Y. Ding, et al., “Broadcast yourself: Understanding YouTube uploaders,” in Proc. ACM SIGCOMM Conf. Internet Meas., 2011, pp. 361–370.

[10]. Q. He, T. Shang, F. Zhuang, and Z. Shi, “Parallel extreme learning machine for regression based on MapReduce,” Neurocomputing, vol. 102, pp. 52–58, 2013.

[11]. A. Basu, S. Shuo, H. Zhou, M. Lim, and G. Huang, “Silicon spiking neurons for hardware implementation of extreme learning machines,” Neurocomputing, vol. 102, pp. 125–134, 2013.