What others have done before us? We have found two studies that used the same datasets as us. The first one created the CMU Viral Videos Dataset which is the "largest public viral video dataset (…) where videos were manually selected by experts including editors from Youtube and the Time Magazine" (Jiang et al., 2018). The team analyzed them and their characteristics to understand how to design a viral video. They made six different and interesting observations, for instance, stating that viral videos have shorter titles and shorter duration in time or that the popularity of the uploader is crucial to viral video views (ibid).

The second research used the Kaggle dataset to study the top 200 Youtube videos trending from November 2017 to March 2018 in the US and the UK. For instance, they looked at views vs trending dates, tags counts, channel subscribers, … making visualizations, and statistical observations out of that information (Niblock, 2018). It is interesting to note that thousands of notebooks of codes revolving around our topic have been made using the Kaggle dataset which was relevant to look at when doing our own research (Jolly, 2018).

Moreover, models such as the HIP (Hawkes Intensity Process) have been created to "quantify the relationship between the popularity of an online item and the external promotion it receives". The research we looked at used HIPs to forecast the popularity of Youtube videos and to identify trends based on the 5-months feed of the most-tweeted videos (Rizoiu et al., 2017).

Our research debunks the Youtube algorithm looking at trending and viral videos. Our study can be qualified as original, firstly, because not any other research we have seen combine both viral and trending videos. Secondly, most studies revolving around our subject usually use Youtube APIs, however, we chose not to. Lastly, since the Youtube algorithm changes frequently there is a need for an up-to-date research.

An API is a set of "programming code that enables data transmission between one software product and another", basically it is data exchange.

References

Jiang, L., Miao, Y., Yang, Y., Lan, Z. and Hauptmann, A., 2018. *Viral Video Style: A Closer Look At Viral Videos On Youtube*. [online] Available at: <http://www.cs.cmu.edu/~lujiang/camera_ready_papers/ICMR2014-Viral.pdf> [Accessed 28 December 2020].

Jolly, M., 2018. *Trending Youtube Video Statistics*. [online] Kaggle.com. Available at: <https://www.kaggle.com/datasnaek/youtube-new> [Accessed 28 December 2020].

Niblock, M., 2018. *Youtube Trending Video - Basic Analysis*. [online] Kaggle.com. Available at: <https://www.kaggle.com/miguelniblock/youtube-trending-video-basic-analysis> [Accessed 28 December 2020].

Rizoiu, M., Xie, L., Sanner, S., Cebrian, M., Yu, H. and Van Hentenryck, P., 2017. *Expecting To Be HIP: Hawkes Intensity Processes For Social Media Popularity*. [online] Available at: <https://arxiv.org/pdf/1602.06033.pdf> [Accessed 28 December 2020].