Determining the Effectiveness of Youtube Videos in Teaching and Learning with Mozdeh Algorithm

Terlumun, I. T.
School of Science Education
Federal College of Education (Technical),
Gombe, Nigeria

Appollm, Y. I., Ibrahim F. J., Mamman, F. S. & Yusuf, A. K.
School of Vocational Education
Federal College of Education (Technical),
Gombe, Nigeria

Ibrahim, A. H.
School of Technical Education
Federal College of Education (Technical),
Gombe, Nigeria

Abstract
This paper is carried out to investigate Numberphile, a mathematics YouTube video channel and determine the academic effect of the video on the viewers. The video was uploaded on the YouTube social Network by experts for over five years with the intention to either upgrade the already existing knowledge or give fresh knowledge to the learner. Much comment is made as feedback by the viewers over a long period of time expressing their feelings. These comments are difficult to classify ordinarily without an algorithm since every word or statement has two dimensions, positive and negative. A new algorithm, Mozdeh Window Free Software is used to determine the sensitivity and strength of the sentiments from emotions and or feelings in the extracted comments on the video. It is able to predict positive emotion with Average and 95% confidence intervals: Pos 3.0436 (3.0342, 3.0530) accuracy and negative emotion with Neg 1.4476 (1.4310, 1.4643) accuracy, both based upon strength scales of 1-5.

Keywords: *Numberphile, YouTube, Mozdeh, Sentiment, Comments*

Introduction

YouTube is one of the most successful social website according to Cheng, Dale, and Liu, (2008), which is providing short video and sharing services since its establishment in early 2005 for the new generation. Its impact on Internet traffic these days is noticeable though its scalability has suffered a great deal (Cheng, Dale & Liu, 2008). A lot of comments numbering in hundreds of thousands are posted every day on Youtube videos in either support or criticism. Quite often than not emotions have been frequently significant in these texts for expressing feelings thereby displaying social support or as part of online arguments (Thelwall et al, 2010). YouTube videos indicate to have very strong correlations with each other and allow privileges for building increased quality of techniques and service enhancement. It has become overtime the quintessential Web 2.0 site due to user generated content and abundance tagging (Gill et al, 2007).

YouTube has no limit on the total number of videos that a user could upload rather it offers a wide range of basic community functions such as the privilege to connect to other users across the globe as friends and also guaranteed URLs and HTML code which makes embedding into other websites easy (Burgess & Green, 2013). They went ahead to substantiate that the site gave an extremely straightforward incorporated interface which clients could transfer, distribute and see spilling recordings without abnormal state of specialized information and inside the innovative limitations of standard program programming and relative humble bandwidth (Burgess and Green, 2013). YouTube also provides an analytic as a tool for its users thereby making it easy for them to track all received and viewed videos and knows which countries contribute the most.

The academic research has tried to understand the tentacles that shape the YouTube as a participation space which go beyond traditional oppositions of romanticizing users as the Web 2.0 revolution trend is seen to be corrupting the young across the globe (Keen, 2011; Muller, 2009). There has been a popular discourse about YouTube participation which has been reshaped to include popular discourse where participants engage on questions of knowledge and skills (Muller, 2009). There is a discourse about this social website YouTube which engages in making videos for academics with teaching skills which is beyond the suspicions of the conservative cultural elites (Muller, 2009; Thoene, 2012; Sue, 2012). People are privileged to update their knowledge in various fields by utilizing the videos on YouTube sites.

There has been evidence that digital intervention has positive effects on the development of expertise in algebra among mathematics teenager students (Bokhove & Drijvers, 2012). One could not deny the fact that YouTube is a good tool of digital presentation. Gromik, (2012) in his case study of cell phone recording feature as a language learning tool substantiated that technological advancement requires educators to understand the benefits of short videos making for enhancing learning in classrooms hence the YouTube Web 2.0.

This medium has made available several social tools for an effective community interaction (Thoene, 2012), and has also made it possible for users to make comments on the videos and as well rate the comments and help in filtering for efficient and relevant opinions (Thelwall, Sud, & Vis, 2012; Siersdorfer et al, 2010). This helps a great deal in obtaining the potential and interesting information to mine and determine users' implicit understanding for the interest of the community (Siersdorfer et al., 2010.). Although YouTube is used predominantly for entertainment, lot academics have ventured into video recording to publicize their scholarly related activities online via YouTube (Kousha, Thelwall, & Abdoli, 2013). Videos of topics relevant to learning are on the increase.

YouTube in Academic Education

YouTube usage for presentation of lectures has been adopted by individual teachers who are computer literate as many academics now use it for recording and disseminating course lectures. These videos are very useful because even students who miss such classes or when teachers are away (Pasquali, 2007 in Kousha, Thelwall, & Abdoli, 2012; Burke, & Snyder, 2008), stand to benefit from the lectures. Besides, most videos can also be viewed by other potential users across the globe.

Although YouTube videos have shown much potential uses to support teaching in the academia, there still exist doubts about the accuracy of these YouTube videos for academic teaching (Freeman & Chapman, 2007; Hossler & Conroy, 2008; Pandey, Patni, Singh, Sood, & Singh, 2010 in Kousha, Thelwall, & Abdoli, 2012), perceive there might be some inaccurate information especially in medicine and public health. Hence the guard against unauthorized usage of such sources.

It was discovered Kousha, Thelwall, & Abdoli, 2012), that the YouTube channel of the University of California, Berkeley (youtube.com/user/UCBerkeley) contains over 3,000 lecture videos on various subject areas which has fascinated more than 13 million views. In the same vain, (Burke, & Snyder, 2008) also suggests that professors of college health education should endeavour to capture the Web 2.0 generation of learners by use of innovative video technology resources such as YouTube which can be meshed to provide relevant and targeted information to add to college course content. It might be an enrichment of curriculum and learning itself which could be very useful to learners.

Several educators maintain that having learners create content as part of their course requirements it would be a necessary element to stimulate learning (Burker, & Snyder, 2008). YouTube encourages content creation and has the prospects of exposing learners to new insights and skills, as well as engage students in online communities. It also creates content and allows students to progress by having a profound assimilation of the material and gives opportunity for experiential learning of the technology used and the content (Burker, & Snyder, 2008; Hall, 2010 in Clifton, & Mann, 2011).

It has been discovered also (Brook, 2011; Muniandy & Veloo, 2011; Cayar, 2011 in Kousha, Thelwall, & Abdoli, 2012) that the teaching and learning of Mathematics, Technical and Language requires personal experiences and latent uses of YouTube videos in academic education which have been discussed in women's studies (Hoskins, 2009), and history (Rees, 2008) in the arts and humanities. Such issues were also discussed in engineering (Fernandez et al., 2011; Kaw & Garapati, 2011), in agriculture (Settle et al., 2011), in computer science (Carlisle, 2010), in chemistry (Franz, 2012), in sport sciences (Burden & Parker, 2008), in dentistry (Knösel, Jung, & Bleckmann, 2011), and nursing education (Clifton & Mann, 2011), which disclosed a broad interest within science and medicine (Kousha, Thelwall, & Abdoli, 2012).

Reference administrators store scholarly references and may enable clients to distribute or share their references or create reference records (Thelwall, & Kousha, 2016). Online reference directors that permit reference sharing (maybe basically for diary articles: Borrego and Sear, 2012), incorporate Mendeley, Bibsonomy (Hotho, Jäschke, Schmitz, and Stumme, 2006, 2007; Mitzlaff, Benz, Stumme, and Hotho, 2010), CiteULike (Bogers and Bosch, 2008), and Zotero (Ritterbusha, 2007). Reference sharing can happen by going to similar authors' reference records or through social labelling (Zanardi and Capra, 2008; Lee and Schleyer, 2012). At the season of composing, Mendeley (now possessed by Elsevier) enabled clients to list their own particular articles on their profile despite the fact that the site is by all accounts predominantly centred around sharing reference records as opposed to researchers publicizing their findings (Thelwall, & Kousha, 2014). Some reference chiefs, for example, RefWorks (Hristovaa, 2012), likewise permit reference sharing, in spite of the fact that this is not their essential capacity (Thelwall, & Kousha, 2014).

Mendeley was made to utilize synergistic refining to enable clients to discover references by associating with others that are comparable (Henning and Reichelt, 2008 in Thelwall, & Kousha, 2014), and there is CiteULike proof that this works (Bogers and Bosch, 2008 in Thelwall, & Kousha, 2014). Mendeley readership measurements have appeared to relate with scholastic references (Li, Thelwall, and Giustini, 2012 in Thelwall, & Kousha, 2014), affirming the insightful idea of the site. One little scale contemplates recommends that

Mendeley may file the lion's share of articles of scholastics in a few ranges (Bar-Ilan et al., 2012 in Thelwall, & Kousha, 2014).

Inside the science estimation field, scientometrics, new wellsprings of data, such as YouTube, can prompt the improvement of new pointers for part of insightful execution (Thelwall, Kousha, Weller, & Puschmannn, 2012). These new markers should be assessed before use to examine that they are sensible and to get to their confinements. The checking can happen in an assortment of ways, yet a typical fundamental method is to correspond to the new markers of better known esteem (Moed, 2005b in Thelwall, Kousha, Weller, & Puschmannn, 2012). For YouTube, neither approach appears to be conceivable since, as can be seen from the outcomes, recordings are moderately divergent as far as fields, influencing reliable companion to survey judgements hard to get (Thelwall, Kousha, Weller, & Puschmannn, 2012).

The purpose of this paper is to gather data on Numperphile YouTube video filter and test the comments with Mozdeh Window Free Software. The filtered data will be processed and used for analysing the sentiment strength in determining whether the academic video channel (Thelwall, 2017). It also determines the impact the YouTube video on the viewers with regards to their comments. It focuses on strength of gender's positive and negative sentiments using matching keywords as well as sentiment and association mining from search matches of the filtered comments across the years that the videos have existed (Thelwall, 2017; Umesh Arya, 2016c).

Exploring YouTube's Application Program Interface (API)

The YouTube Data API enables you to join works regularly executed on the YouTube site into your own site or applications. A run-down recognizes the different sorts of assets that you can recover utilizing the API. The API likewise underpins strategies to embed, refresh or erase a significant number of these assets (Google Developer, 2017). Every blend of style and access strategy has its own particular advantages and suggestions. The style of the API will affect its ease of use crosswise over different expending engineering/dialect models (standards) and may impact the general style of the devouring application (Frye, 2015). The technique for access may make levels of manufacture time and organization conditions and openings that must be looked at when as an expending application is choosing the best arrangement (or on the off chance that it can be utilized by any stretch of the imagination), (Frye, 2015). The Application Program Interface (API) is the information sharing technology where YouTube interact with to download needed post (Thelwall, 2017). Every blend of style and access strategy has its own particular advantages and suggestions. The style of the API will affect its ease of use crosswise over different expending engineering/dialect models (standards) and may impact the general style of the devouring application (Frye, 2015). The technique for access may make levels of manufacture time and organization conditions and openings that must be looked at when as an expending application is choosing the best arrangement (or on the off chance that it can be utilized by any stretch of the imagination), (Frye, 2015).

In obtaining a developer key for YouTube Data API v3, one needs to have a Google Account to be able to access developers console for the request of an API key to register an application. By this, the intended programmer will create a project to obtain authorization credentials for the application to be submitted for API requests making sure that the YouTube Data API is that which your application is registered to use. The process will guide you to finish as designed to be followed by YouTube > Data API (Thelwall, 2016;).

One solution has been to innovate and build technology helps and assist storage arrays to be situated on the network directly where disk accesses may be made directly rather than through the server's (SCSI) Small Computer System Interface (Bowen et al, 2008). He has substantiated that the storage area network (SAN) representing capacity range demonstrate places Storage all alone devoted system, expelling information stockpiling from both the server-to-circle SCSI transport and the principle client arrange. This devoted system most generally utilizes Fiber Channel innovation, a flexible fast transport (Bowen et al, 2008).

A medical research was carried out a few years ago to develop a technique that will aid the extraction of unfolding drugs abuse signs and symptoms from YouTube video comments about marijuana using the YouTube's API and results were generated by using a Bernoulli Naive Bayes classifier (McKenzie, Park, & Manini, 2013). Song, & Gruzd, (2017) in their research on examining Sentiments and popularity of pro- anti-vaccination videos on YouTube made use of YouTube's API where they were able to retrieve as much videos as needed using some keywords and thereafter analysis generated. Every day, over 150,000 new videos are added mostly generated by users but much is not known who the YouTube users for whom they upload these videos. Courtois, Mechant, and De Marez (2011) in the second face of their survey decided to use YouTube's Application Programming Interface which they christened 'API; a gateway to interact with the platform' to select uploaders who would participate in an online interview. Their disclosure that teenage uploaders after a latent growth analysis was performed indicated public expectancy corresponded fairly with the feedback received (Courtois, Mechant, and De Marez, 2011). Thelwall, (2017) in his book titled "Social Web Research with Mozdeh" confirm that the API does not retrieve all and old tweets as well as spam-like tweets.

Exploring for Texts Comments on YouTube

The participatory Web technologies, "Web 2.0" have prompted the people's interest in the keen use of multimedia on social networks by expanding the means of persuasive messages written out to users (Walther et al, 2010). YouTube comments are very vital in carrying out studies concerning users where their individual opinion is retrieved in the analysis of such studies (Thelwall, 2017; Mike Thelwall, 2014). Although the openness in the participation of users of these technologies for purposes of analysis may be subverted as contradictory sentiments posts, there is always an outcome from it, there is no doubt that they go a long way in achieving the intentions of researchers.

Recent Web technologies encourage users to post comments about content where commenting sometimes takes the form of reactive exchanges, but in some cases, participative technologies encourages interactive exchanges such that despite their carefree nature evolve to dialogue (Rafaeli, 1988 in Walther et al, 2010).

On so many occasions the YouTube reveals a large amount of community feedback through comments for published videos and also through Meta ratings for these comments where they are retrieved and filtered to determine sentiments of such users. The influence of sentiment expressed in comments on the ratings for these comments using the SentiStrenght (Thelwall, 2017; Sierasdorfer et al, 2010), a lexical WordNet-based resource containing sentiment annotations.

The creative use of punctuation otherwise known as emoticons when text messaging is the most pronounced verbal features in the social web has become a custom for abbreviation; the initialism which (Thelwall, 2017; Thelwall et al, 2010) referred to as slang lookup table in his SentiStrength data . This is the use of initial letters meant to represent the entire words, like

tnx for *thanks*; the omission of nonessential letters especially homophones, like *b4n* for *bye for now*; and the substitution of vowels like *cul* for *see you later*. This unique style of texting is, as a result of trying to keep it short, that is keeping messages within the 160-character limit (Crystal, 2008, pp. 5-6 in Jones, & Schieffelin, 2009).
There are many web platforms that are used to share non-textual content, such as videos, images, and animations with which users are enabled to add comments from among them (Serbanoiu, & Rebedea, 2013; Thelwall, Wilkinson, & Uppal, 2010). The most accepted is the YouTube with massive videos uploaded by its users on a regular basis followed by millions of comments. Most of the videos have just a handful of text comments while some have comments amounting to millions without a clear filtering mechanism. Thelwall, (2017) has recently developed free window software christened MOZDEH for news analysis.

Collecting Data with Mozdeh
Researchers have been eased the stress of trying to collect and analyse comments on YouTube and Twitter with the invention of Mozdeh free Windows software. Thelwall, (2017), described mozdeh to mean "Good News" and narrated that it is derived from the Mozdeh origins in Persia. Before now, mozdeh was incorporated in Webometric Analyst in carrying out transactional process of analysis (Thelwall, 2016; Umesh Aryah, 2016c), where plain texts file only was used in collecting and analysing data. The jinx has been loosened since big data can be gathered and analysed on YouTube and twitter by mozdeh which is free at Mozdeh website (http://mozdeh.wlv.ac.uk).

Mozdeh can gather tweets or texts comments by matching one or more queries which are listed in the data collection screen once it is started. It carries out process by submitting them to YouTube and or twitter, via its Application Programming Interface (API) and then save all matches returned (Thelwall, 2017), repeating continuously checking for new content. The magic software as I call it can collect data non-stop for years provided your computer and power are not turned off.

Gathering both tweets and texts on YouTube involves the use of matching keywords (Thelwall, 2017), from a set of users is now done with Mozdeh straight away with the support of YouTube and twitter for which it is developed for. This free software can gather texts on social web with approval from an Application Programming Interface (API) which serves as the information sharing technology for downloading data with Mozdeh (Thelwall, 2017). Mozdeh has the capacity to gather texts or tweets matching one and or even more queries through the Mozdeh data collection screen which will be displayed later on.
Twitter permits the collection of two types of texts (data) with Mozdeh for free provided the tweets matching the queries are new, but all older tweets must have to come from specific and known users (Thelwall, 2017).The emergence of the social media around the middle 2000s has opened up new paths to studying both social and cultural procedures and dynamics.in new ways (Manovich, 2011). Images and the videos are seen and comments made on them after monitoring the discussion engaged in, read blogs and tweets, listen to tracks lists without the owners' permission.

Mozdeh can accumulate social web messages straightforwardly from sites that offer them, for example, twitter and youtube, or can import messages that have been assembled from different ways (Thelwall, 2017). Both twitter and YouTube as of now share their information through an Application Programming Interface (API) which is a data sharing innovation that Mozdeh associates with to download posts (Thelwall, 2017).
There is always the gathering aspect of data in any given research hence for the benefit of this

piece of work Mozdeh free Window software is made use of. (Umesh Arya, 2016b).

If we consider Google search for example, its algorithms analyse billions of web pages, Excel spreadsheets, Word documents, Flash files, plain text files, plus PDF, Facebook, Twitter, and YouTube content (Manovich, 2011), exposed that Google does not allow users to analyse patterns directly in all of its data the way Google Insights searches queries and Google Ngram Viewer searches digitalized books. (Manovich, 2011) describes a hypothetical scenario of working with a million YouTube documentaries–style videos which computers could be used to investigate quickly massive visual data sets and thereafter chose the objects for a closer physical analysis.

There are platform-specific sampling muddles such that the biggest-volume source of public Twitter data which is used by lot of researchers across the globe of the overall platform's data is not accurately represented (Morstatter, Pfeffer, & Liu, 2014 in Derek, & Pfeffer, 2014). More so, it would not be wrong to say researchers are left in dilemma about accurate data since social media providers may change the sampling and/or filtering of their data streams which affects the algorithms and processes that govern these public data (Derek, & Pfeffer, 2014). Academic attempts to represents aspects of the behaviour of such ownership-syndrome systems can provide features needed to begin bias publishing so long as the bureaucracy and procedures governing the release of these public data remains complex (Derek, & Pfeffer, 2014).

Analysing the YouTube Comments
YouTube has been discovered to be a social phenomenon where individuals are permitted to create and share media content with many other users around the globe (Thelwall, 2014), which serves as a valuable source of information for diverse of mainstream and as it does a huge disposition of videos on various topics.

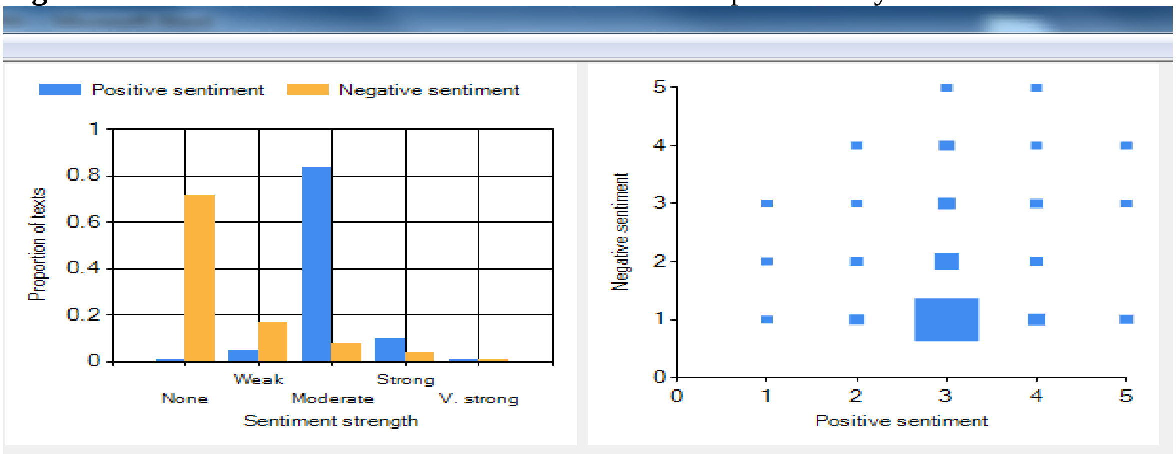
Overtime some software and apparatus have been in place for analysing YouTube comments. Before the new invention of Mozdeh, Thelwall, (2014), had previously made use of the then new software application called Webometric Analyst which has capacity as developed by the author to download comments on videos and public information about commenting users such as their age, sex and demography. These networks permit researchers who narrated logically the debate surrounding the responses of the audience.

Mozdeh carries out analysis on sentiment to detect and codify subjective content and estimate the strength of positive and negative sentiment in the texts using the SentiStrength (Thelwall, 2017; Mike Thelwall, 2013c), which also detects sentiment by use of lexicon of sentiment. This assigns two scores to each text between (1) meaning no positive sentiment and (5) meaning very strong positive sentiment and the same applies to the negative sentiment which has (-1) no negative sentiment and (-5) very strong negative sentiment (Thelwall, 2017; Pedrycz & Chen, 2016; Thelwall, Buckley, Paltoglou, 2012). The numbers [-1, 1] show that there is no negative and positive sentiment while the numbers -3, 5 specify a moderate negative sentiment and a very strong positive sentiment. Sentiment is calculated based on P, N and O where every P and N represents the counts of positive and negative codified sentiment words while O is the count of all words other than positive and negative such that the total number of words equals $P+N+O$. The Absolute Proportional Difference with bounds: [0, 1] is calculated as:

$\text{sentiment} = \frac{(P-N)}{(P+N+O)}$. The Relative Proportional Difference with bounds: [-1, 1] is calculated as: $\frac{(P-N)}{(P+N)}$ while the Logit scale, with bounds: [-infinity, +infinity] $\text{Sentiment} = \log(P+0.5) - \log(N+0.5)$ and this tends to have smoothest properties and is symmetric around zero. Hence the 0.5 is smoother in order to prevent log (0). (Lowe et al., 2011).

The output of Mozdeh after calculation is displayed showing the classification of all texts in the project for both the positive and negative sentiment and giving feedback of scores and texts on the screen showing the bar chart and bubble chart (Thelwall, 2017; Mike Thelwall, 2014).

Figure 1: The Bar chart and Bubble chart for Sentiment produced by Mozdeh search.



Mozdeh Free Software also does the sorting and filtering of sentiment by selecting one of the four options for positive and negative sentiment either in ascending or descending order based on the type of sentiment that it is detailed to calculate (Thelwall, 2017). The overall result is displayed for the tasks that are requested of it in averages along with a 95% confidence interval.

Sentiment analyis results

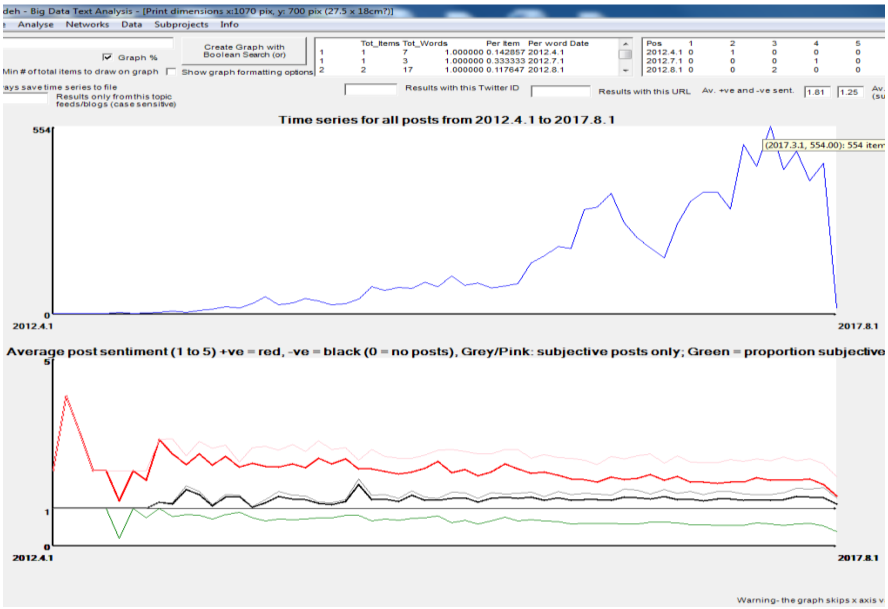
Score	**Pos**	**Neg**
1	0.87%	71.26%
2	4.96%	16.89%
3	84.04%	8.02%
4	9.21%	3.46%
5	0.92%	0.36%

Average and 95% confidence intervals:
Pos 3.0436 (3.0342, 3.0530)
Neg 1.4476 (1.4310, 1.4643)
Av pos - Av neg: 1.5960
To see bootstrap sentiment confidence intervals (slow) recalculate after ticking: Analyse menu, Calculate bootstrap sentiment confidence intervals

Conclusion

The results as presented in this chapter are derived from the data collected and processed by Mozdeh free software. The participants are commenters of the identified YouTube video; the Numberphiles. These commenters are people who have taken their time to watch the distinct video and perhaps have benefited a lot from. Those comments have been retrieved by Mozdeh and critically analysed. It gathered data in form of comments made on the video with matching keyword from users. All this is incorporated data gathering, topic content exploration, topic content description, sentiment analysis, gender difference identification, time series analysis, association mining and of course networks analysis (Thelwall, 2017; Thelwall, 2016; Mike Thelwall, 2014). All the points mentioned have been carefully performed and results provided.

Figure 2: Time Series Table for Posts.



This table has one major and one minor spikes as well as a trough though it's a cyclic trend graph. The outputs for average positive and negative sentiments are 1.81 and 1.25 while that of subjective average positive and negative sentiments are 2.34 and 1.41 respectively.

It has been confirmed (Thelwall, 2017; Thelwall, 2016) that the social web has tremendously aided lots of academics and non-academics alike in voicing out their feelings and ideas as publications and social interaction on a more larger scale as hitherto done. The social web has created a level playing ground for all and sundry. A lot of activities are carried out on social networks as it is now considered to be the most convenient medium of communication. Several options are available for people to make selections from such as the YouTube and a host of others.

References

Bokhove, C. and Drijvers, P. (2012) Effects of a digital intervention on the development of algebraic expertise. *Computers & Education* [online], **58**(1), pp. 197-208. [Accessed on 2 August 2017].

Bowen, J.F., Mathkar, A.R., Mathur, R., Syed, S.M., Weimer, T.W., Bennett, J.E., Braganza, C.W. and Dwivedi, T. (2008) *System and method for an access layer application programming interface for managing heterogeneous components of a storage area network* [online]. [Accessed 25 August 2017].

Burgess, J. and Green, J. (2013) *YouTube: Online video and participatory culture.* [online] John Wiley & Sons.

Burke, S.C. and Snyder, S.L. (2008) YouTube: An Innovative Learning Resource for College Health Education Courses. *International Electronic Journal of Health Education* [online]**, 11**pp. 39-46 .[Accessed 25 August 2017].

Clifton, A. and Mann, C. (2011) Can YouTube enhance student nurse learning? *Nurse education today* [online]**, 31**(4), pp. 311-313 .[Accessed 25 August 2017].

Cheng,X., Dale,C. and Liu,J. (2008) Statistics and social network of youtube videos *Quality of Service, 2008. IWQoS 2008.16th International Workshop on.*[online]. IEEE, pp.229-238.

Courtois, C., Mechant, P., and De Marez, L. (2011) Teenage Uploaders on YouTube: Networked Public Experiences, Online Feedback Preferences, and Received On-Platform Feedback, Cyber psychology, *Behaviour, and Social Networking*, **14** (5), pp. 315-322.

Freeman, B. and Chapman, S. (2007) Is "YouTube" telling or selling you something? Tobacco content on the YouTube video-sharing website.*Tobacco control* [online]**, 16**(3), pp. 207-210 .[Accessed on 2 August 2017].

Gill,P., Arlitt,M., Li,Z. and Mahanti,A. (2007) Youtube traffic characterization: a view from the edge *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement.* [online]. ACM, pp.15-28.[Accessed 25 August 2017].

GOOGLE DEVELOPERS, 2017.YouTube Data API [Online].Google. Available from: http://developers.google.com/youtube/v3. [Accessed 9 September 2017]

Gromik, N.A. (2012) Cell phone video recording feature as a language learning tool: A case study. *Computers & Education* [online]**, 58**(1), pp. 223-230. [Accessed 2 September 2017].

Jones, E.M and Schieffelin, B.B. (2009)Talking Text and Talkingc Back: “My BFF Jill” from Boob Tube to YouTube, *Journal of Computer-Mediated Communication*/Volume 14, Issue 4, https://doi.org/10.1111/j.1083-6101.2009.01481.x

Keen, A. (2011) *The Cult of the Amateur: How Blogs, MySpace, YouTube and the Rest of Today's User Generated Media are Killing Our Culture and.* [online] Nicholas Brealey Publishing. [Accessed 2 September 2017].

Kousha, K., Thelwall, M. and Abdoli, M. (2012) The Role of Online Videos in Research Communication: A Content Analysis of YouTube Videos Cited in Academic Publications, *Journal of the Association of Information Science and Technology* **63** (9), pp. 1710-1727.

Lowe, W., Benoit, K., Mikhaylov, S., and Laver, M. (2011), Scaling Policy Preferences from Coded Political Texts, Legislative Studies Quarterly 36 (1), pp. 123-155.

Manovich, L. (2011) Trending: The promises and the challenges of big social data.*Debates in the digital humanities* [online]**, 2**pp. 460-475 .[Accessed 1 September 2017].

McKenzie,A. (2013) An Analysis of YouTube Comments on Drug Health Effects *Medicine 2.0 Conference.* [online]. JMIR Publications Inc., Toronto, Canada, .[Accessed on 2 August 2017].

Müller, E. (2009) Where Quality Matters: Discourses on the Art of Video Making for YouTube. [online]. [Accessed 2 September 2017].

Mike Thelwall, 2013c. Expressing and Not Expressing Sentiment in Social Media [Online video]. Available at: http://www.youtube.com/watch?v=JE-5Nd6dzGU [Accessed 15 August, 2017].

Mike Thelwall, 2014. Sentiment Strength Detection For The Social Web. [Online video].Available at:http://www.youtube.com./watch?v=5hx8qta-XKg [Accessed 9 September 2017].

Pedrycz, W. and Chen, S. (2016) *Sentiment analysis and ontology engineering: an environment of computational intelligence.*[online] Springer.[Accessed 16 August 2017].

Ruths, D., and Pfeffer, J., (2014) Social Media for Large Studies of Behaviour, *(American Association for the Advancement of Science),***346** pp.1063-1064.

Siersdorfer,S., Chelaru,S., Nejdl,W. and San Pedro,J. (2010) How useful are your comments?: analyzing and predicting youtube comments and comment ratings *Proceedings of the 19th international conference on World wide web.* [online]. ACM, pp.891-900.[Accessed 2 September 2017].

Song,M. Y. and Gruzd,A. (2017) Examining Sentiments and Popularity of Pro-and Anti-Vaccination Videos on YouTube *Proceedings of the 8th International Conference on Social Media & Society.* [online]. ACM, pp.17.[Accessed 16 August 2017].

Serbanoiu, A and Rebedea, T., (2013)Relevance-based Ranking of Video Comments on YouTube, *Control Systems and Computer Science (CSCS), 2013 19th International Conference* on, pp. 225-231

Thelwall, M., Kousha, K., Weller, K., Puschmann, C. (2012) Assessing the Impact of Online Academic Videos, *Social Information Research*, pp. 195-213

Thelwall, M. and Kousha, K. (2014) Academia. Edu: Social Network or Academic Network?, **65** (4), 721-731.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. and Kappas, A. (2010) Sentiment strength detection in short informal text. *Journal of the Association for Information Science and Technology* [online]**, 61**(12), pp. 2544-2558 .[Accessed 16 August 2017].

Thelwall, M., Sud, P. and Vis, F. (2012) Commenting on YouTube videos: From Guatemalan rock to el big bang. *Journal of the Association for Information Science and Technology* [online]**, 63**(3), pp. 616-629 .[Accessed 16 August 2017].

Thelwall, M. (2017) The Heart and Soul of the Web? Sentiment Strength Detection in the Social Web with SentiStrength.*in Cyberemotions.* Springer,.[Accessed 6 August 2017].

Thoene, W.S. (2012) The impact of social networking sites on college students' consumption patterns. [online]. [Accessed 10 August 2017].

Umesh Arya, 2016c. Sentiment Analysis of Words in Mozdeh.[Online video]. Available at: http://www.youtube.com/watch?v=VDPyvgtLV-0 [Accessed 5 August, 2017].

Walther, J.B.,DeAndrea, D., Kim, J., Anthony, C. A. (2010) The Influence of Online Comments on Perceptions of Antimarijuana Public Service Announcements on YouTube. *Human Communication Research* **36** (4), pp.469-492