# *Sentysis*: A Sentiment Analysis Algorithm for YouTube Comments


**ABSTRACT**

YouTube is one of the most popular non-textual media platforms to spread information and ideas on the Internet. Viewers of these media can post comments on a specific video. There has been a multitude of research about these user comments, with the most popular being sentiment analysis. This paper will focus on using both methods of analysis on news or political YouTube videos and their respective comments. The results will be indicative of how divisive a political topic may be and how knowledgeable the viewer may be about the topic. However, this paper will not be able to prove any causal claims; instead, the research can only be used to establish a correlation between viewers' attitudes and a news topic.



Keywords: climate change, natural language processing, sentiment analysis, YouTube.


## I. INTRODUCTION

In the modern age of technology, there is a lot of information being spread on the Internet; in fact, according to Roser et. al [1], 3.4 billion people (as of 2016) have used the Internet, meaning that there are billions of ideas being exchanged on the Web. One of the most notable platforms for these ideas is YouTube, a site that allows anyone to upload a video about nearly anything and watch others' videos. According to SimilarWeb [2], the site gets around thirty billion visits monthly, and it is the second most popular website in the world.

Therefore, a lot of information can be spread, and the public is more informed than it has ever been. On the other hand, it can be argued the public is now more misinformed than it has ever been. People may use sources to confirm their misinformed opinions; this process is known as confirmation bias. For example, Yiannakoulias et al. [3] state that anti-vaccine content on YouTube tends to be more "liked" than pro-vaccine content on YouTube. This means that there are more people searching for anti-vaccine content that confirms their minority opinion, rather than the majority of people that have accepted that vaccines are safe. This problem can also be in user comments on YouTube, though. Uryupina et al. [4] state that in a video concerning a well-known product, like the iPad, experts' comments will use the correct terminology of that field, and include evidence/sources. A lesser-informed user, though, may seem more emotional and rash in his/her comment.

Uryupina et al. used a specific method in their research study called sentiment analysis; sentiment analysis is a method that analyzes the sentiment of a given text and returns its mood and the strength of its mood. The analysis, of course, must have a scale. For explanatory purposes, it could range from negative one to positive one, where the negative one would represent the strongest negative sentiment, and the positive one represents the strongest positive sentiment. If this were to be applied to a sentence that reads, “I am glad to have bought this product”, the sentence would most likely receive a positive one. However, sentiment analysis is much more complicated; the scales can often be inaccurate if done without the proper criteria.

Before further elaborating on the measurements or quantitative results of sentiment analysis, it must be understood that most sentiment analysis is not done manually; rather, it is usually automated through computer algorithms and programs. These algorithms were, however, built and coded by engineers and computer scientists. For example, Dr. Alexandra Balahur [5], an Italian computer scientist who has published numerous papers on emotion detection and sentiment analysis, published a paper and created an efficient algorithm to analyze the sentiment of tweets.

The measurement of the sentiment and polarity of a text may vary greatly depending on what scale is used. First, the scale must be defined using a dictionary; a dictionary is a list of words or phrases that already have a set polarity and sentiment assigned. “Good” may have a fairly strong positive score, while “sucks” may have a fairly negative score on the scale. These dictionaries are often defined from a multitude of institutions or third-parties. One example is the Harvard-IV dictionary directory [6], where there are two dictionaries for negative and positive words or phrases. Even within this one directory, though, there are many dictionaries that interpret words and phrases through a different lens. There are dictionaries for religious, legal, business, or academic settings. Some even are dictionaries dedicated toward a certain feeling: failure, virtue, arousal, or pain. Because of these many dictionaries, it can be concluded that the discourse community does not have a standardized dictionary for every setting and every feeling; therefore, sentiment analysis will always have some margin of error. For instance, if a commercial (marketing or business uses) dictionary was used for sentiment analysis on ratings for a thriller movie, the word unpredictable would give a negative score to the overall sentiment of the ratings; however, in this circumstance, unpredictable is likely to have a positive connotation. Because of such flaws, sentiment analysis is not by any means perfect; though, it can give a detailed analysis of the emotions of the audience.

Sentiment analysis is often utilized in studies concerning news articles and share prices of stocks. An example of this is shown by Bollen et. al [6], in a well-known study. The study analyzed the sentiment of tweets, and using previous data, the algorithm predicted the movement of a stock with 87.6% accuracy. It goes to show that sentiment analysis can be used as a practical method of prediction. As there are some flaws, sentiment analysis is also a topic of research itself; with these practical applications, the discourse community strives to perfect this method of analysis.

Though, Bollen et. al's study represents much of the research done with sentiment analysis. Some research is also conducted to hone the method of sentiment analysis, as shown by Balahur. Sentiment analysis, although complicated, is far more developed and discussed than, when compared with, information quality analysis on YouTube comments. The level of information quality, or as Serbanoiu et. al [8] state, relevance-based ranking, is a helpful analytical reference that will be used to sanitize data in this research. Serbanoiu et. al defines this ranking as a means to filter between irrelevant and relevant comments, based on the YouTube video's subject. Similarly, this paper will aim to filter irrelevant comments. In this paper, YouTube will be utilized as a platform because of its vast popularity. All YouTube comments will be filtered first to check whether or not they are spam comments.

This is essential research to many people — including content creators and sociologists alike — as it is useful to know whether the majority of people commenting on news' videos are posting composed and logical, or reckless and emotional comments. If the research shows that the comments are less emotional in a specific genre of videos, then it may be concluded that many comments are generally both relevant and logical. Effectively, this research will attempt to quantify the social climate of the YouTube community through testing different variables and sentiment analysis; additionally, it will test the correlations between a) controversial topics — climate change, in this case — and a "neutral" sentiment because there are strong polar sentiments and b) who or where the source originates from varies the sentiment either negatively or positively. Therefore, the hypothesis is the research question: does the controversy of the topic and source of the selected YouTube video have a correlation with the viewers' sentiments? How strong or weak is this correlation?

## II. METHODOLOGY

Sentiment analysis itself is a quantitative approach to estimating the sentiment of users. Rambocas and Gama have defined sentiment analysis to be a form of systematic analysis of online expressions. Additionally, the resulting research of this algorithm will not be strong enough to make a causal claim, as there are too many confounding variables — confirmation bias of users and bot comments may be two that can greatly affect the outcome of the analyses (although the proposed methodology will attempt to curb the effects of the latter). However, this algorithm will give a fair foundation to establish more research; if the results show that the majority of users create emotional comments as opposed to logical comments, researchers in sociology, psychology, and even marketing can use this method to craft their own expert conclusions. An expert causal conclusion cannot be made solely from the knowledge in this research alone. Therefore, this study can be viewed as a quantitative correlational study via the Internet, and more specifically, YouTube.

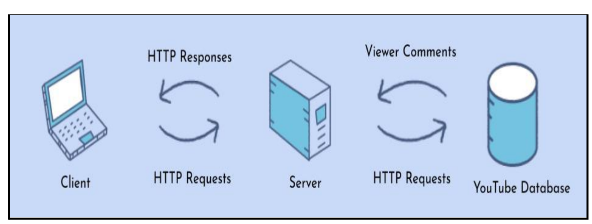
To create a thorough and meaningful research catalog to perform sentiment analysis, there must be a large sample size of comments from a video. The reasoning behind this is derived from the direct relationship between sample size and accuracy of the summary of the users' sentiments; the more comments that can be used to be analyzed, the more accurate the concluding sentiment will become. Additionally, there must be multiple videos on the same topic to sample from; this will allow for a conclusion to be made on at least one topic, which lays the foundation for future research.

*A. Sentysis*

The analysis program, nicknamed *Sentysis*, is formatted as a full-stack web application (Image 1) and can be accessed publicly through a domain (see Appendix, A.1). The user interface (UI), programmed in Angular (see Appendix, A.2), will receive a uniform resource locator (URL) that points to a YouTube video via text input; this will allow for an easy switch between different videos, making the overall research flow more dynamic and efficient. The URL will be sent to the *Sentysis* server, programmed using Flask (see Appendix, A.3). The *Sentysis* server will utilize the public YouTube Data API v3 (see Appendix, A.4) to retrieve the respective comments of the YouTube video. Initially, *Sentysis* would have also requested the video audio from the YouTube database to be able to perform a more accurate analysis of the comments; however, this action is prohibited by the YouTube Terms of Service:

> The following restrictions apply to your use of the Service. You are not allowed to:
>
> 1. access, reproduce, download… any part of the Service or any Content…

Image 1



*B. Data collection and filtering*

Regarding the viewers' comments on the YouTube video, the comments will be chosen sorted by relevance, a parameter provided by the YouTube Data API when requesting data (Image 2). This means that comments cannot be randomly chosen following the guidelines of the simple random sampling method, which would provide the most unbiased and representative sample. Therefore, there may be an unquantifiable bias within each comment sample. However, once the comments are chosen, there will have already been a verification that the comment is a real person. At first, this does not appear to have enough verification. Though, after many trials on sampling with only this filter, each comment sampled was both posted by a real person and relevant to the video. For example, on a YouTube video named "Trump supporters say why they won't take Covid-19 vaccine", this is a comment that was retrieved:

"I'm perfectly healthy, I'm not going to die from it. And you know what, if we do it's our time to go, we believe in god'. Who are these people? Has America really become this stupid? The thing is, these super 'religious' people are selfish, entitled & arrogant. Guess what, if you do get and then give it to someone else and they end up dying ... Well, I guess it was just their time to go also. Absolutely pathetic!"

Image 2

- `relevance` - Comment threads are ordered by relevance.

The author of this comment clearly had a relevant opinion on the topic, and it was not a bot comment. After numerous samples, it can be assumed that YouTube's relevance filter may follow closely to the research of Serbanoiu et. al. Their research found that the following features are considered when deciding whether or not a comment is a spam comment:

- Number of non-ASCII characters
- Number of capital letters
- Number of new-lines
- Number of digits
- Number of trivialities[1]
- Number of words in comment
- Mean word length
- The number of punctuation marks
- Common spam text count

[1] Profanity

Comments with a high number of capital letters, trivialities or low mean word lengths, and word count will be filtered, as these attributes are commonly found in spam comments.

*C. Sentiment Analysis*

After the data collection and filtering have finished, the program will move on to analyzing the sentiment of the comments for each video.

As for the scoring system of sentiment analysis, the program will be similar to the system set by Godbole et. al [10]. Their work was analyzing both blogs and news articles, similar to this paper's subject, news YouTube videos; therefore, it would be fair to have a similar scoring system. Godbole et. al had a simple scoring system for each meaningful word[2]: "*not good* = -1; *good* = +1; *very good* = +2." Though similar, the scale utilized in this research will be from the range of negative one to positive one, where negative or positive represents the polarity of the sentiment, while the numeric value represents the magnitude (zero being the most neutral, while one is the strongest). This is because "*very good*" is a rather vague term; additionally, it will make the research more confusing, as there would not even be a "*very bad*" score. Each comment will be given a score on this scale with no other factors being considered.

[2] Words that add meaning to the overall comment; for example, the word "a" would not be meaningful compared to "fantastic"

Two Python (see Appendix A.5) libraries were considered to calculate the comments' sentiments: Valence Aware Dictionary and sEntiment Reasoner (VADER) [11], and Afinn [12].

VADER, an algorithm publicly licensed by the Massachusetts Institute of Technology and cited by over 2500 developers and researchers, utilizes a complex, yet efficient approach to sentiment analysis. It takes into account the use of negations — words such as but and not — emoticons, punctuation, capitalization, and adverbs, which makes the algorithm more sophisticated than a solely dictionary-based algorithm, like Afinn. The range of sentiment scores corresponds with this study's sentiment score range between negative one and positive one, which means there will be no modification needed to rescale the scores. Additionally, because VADER analyzes the text as a whole rather than word-by-word, there is no need to average the sentiment after using VADER. When testing this algorithm against sample comments for this study, the results were largely accurate, which proved that it should most likely be the primary analyzer algorithm for this research.

Afinn is another popular method of sentiment analysis that uses a simpler wordlist-based approach, or a dictionary. There are a total of 2,477 English words in Afinn's catalog. Words can have values from negative five to positive five, where five is the strongest magnitude of sentiment, and negative or positive denotes the polarity of the sentiment. For example, the word "excellent" has a score of positive three. However, Afinn gives a sum of the sentiment scores in the inputted phrase (in this case, a comment posted by a viewer); therefore, the only solution to utilize this score will be to divide the initial score by the word count and then rescale it to this study's sentiment analysis scale (between negative one and one). The formula for the calculation of the Afinn score in each comment is shown in Formula A. Unlike VADER, though, when testing Afinn's rescaled score against sample comments, the results were largely inaccurate because simple articles such as "a", and "the" were calculated in the score. It made the Afinn score of each comment to skew towards zero, which would be neutral, even when the real sentiment of the comment was rather polar. Filtering out these small words would have made this score more accurate and therefore useable in this project, but there was no real way to do that without major complications that would slow down the *Sentysis* algorithm (i.e. scanning the comment for all the neutral articles or adjectives in the English language, which would take a long time). Therefore, Afinn was omitted from this research study.

On another note, since most YouTube viewers may not post a comment, but they will like or dislike a comment that has been posted by someone else, it is important to also consider the like count of each comment when summarizing the audience's sentiment. *Sentysis* will use the like count of each comment to weigh a comment's sentiment. For instance, if a comment has five hundred likes, five hundred people agree with the comment, and therefore share the same sentiment; on the other hand, a comment with zero likes equates to only one person (the author of the comment) sharing the same sentiment as the comment. Hence, a weighted arithmetic mean, using the like counts as weights, will be calculated to show a more accurate average of the audience's sentiment (refer to Formula A).

Formula A

$$\mu_{sentiment} = \frac{\sum_{n=1}^{50} sentiment_n * likes_n}{\Sigma likes}$$

One metric to measure the sentiment of a sample of thousands of people is not enough; two other metrics were added to contextualize the weighted mean sentiment: the weighted standard deviation from the weighted mean sentiment, and the like-to-dislike ratio on the selected YouTube video.

The weighted standard deviation is simply the standard deviation of the sentiments, taking in account the likes as weights (in the same way the mean sentiment was calculated). To see the mathematical calculation for this weighted standard deviation, refer to Formula B. If this metric is relatively high (greater than the absolute value of the weighted mean sentiment), then it contextualizes the result of the weighted mean sentiment by reporting that there is a high deviation from the weighted mean; therefore, there would be a large variation in the comments. If this metric is low (less than the absolute value of the weighted mean sentiment), then it reports the inverse: most of the data is near the weighted mean sentiment, and therefore there is little variation in the data.

Formula B

$$\sigma_{sentiment} = \sqrt{\frac{\sum_{n=1}^{50} likes_n (sentiment_n - sentiment_{avg})^2}{\Sigma likes}}$$

The like-to-dislike ratio is by far the most colloquial metric, as it is not a standard statistic like mean, standard deviation, or range. It is simply calculating a ratio. Therefore, this final metric should be lightly considered, unless it is overwhelmingly convincing in the direction of a negative or positive sentiment. Though, instead of collecting data from each comment — what has been done with weighted mean sentiment and weighted standard deviation — the total likes and dislikes of the selected video are gathered, and the ratio is then calculated using Formula C. This metric is used to either support or reject the calculations of the previous two metrics; since it is the last metric considered, it is also the last resort when concluding whether or not the sentiment of a video is positive or negative. For example, if the like count is over 100, it can be safely concluded the sentiment is positive if the weighted mean sentiment and standard deviation do not conclude that the sentiment is negative. Inversely, if the like count is less than 10, it can be safely concluded that the sentiment is negative if the weighted mean sentiment and standard deviation do not conclude that the sentiment is positive. These ranges come from years of prior knowledge about YouTube videos.

Formula C

$$Like: Dislike = \frac{\Sigma likes}{\Sigma dislikes}$$

## III. RESULTS

Though the algorithm used in this research can be used for any topic from any YouTube video on the platform, the data collected in this research will be on the specific topic of climate change — an issue that is commonly and passionately debated, with or without scientific evidence. It was chosen for this research because it is a controversial topic, which should have a wide range of sentiments within the comment section of each video; therefore, it should correspond with the possible relationship between a highly controversial topic and a high likelihood of neutrality within the sentiments collected.

Because each video selected was about climate change, the video topic is held constant, so it is not a confounding variable when considering whether or not the source of the video correlates with the viewers' sentiments. However, it is important to consider that all the viewers from Video A, for example, are not the same viewers that watched Video B; because YouTube is such a popular platform, each viewer will have a different sentiment on any topic because they are different — essentially, each person is unique and has their own opinion. A limitation of this research is not being able to hold this variable constant because this algorithm cannot keep people from not watching or force people to watch a certain video.

Additionally, because this research aims to see the association between source and viewer sentiment, two different factors were used to select videos: popularity and location of the source. For example, if a video has millions of views, then it will be considered mainstream as many people have watched the video, and vice versa — a video that only has a few thousand views will be considered fairly unpopular relative to all other videos. Location of the sources will only be differentiated by the country it is from, with a focus on Western versus Eastern.

The first video, titled "The Heat: Climate Change Challenges for 2021" posted by China Global Television Network (CGTN) America [13], was selected because a) it is about climate change, and b) the source is from the American division of a Chinese media entity. Because the source is not mainstream and foreign, it is possible this correlates with the negative sentiment. The results are displayed in Table A. After manually viewing the video's comments, it can be confirmed that the comment base displayed a negative sentiment toward climate change — not necessarily meaning they oppose acting on climate change, but they may oppose certain guidelines or political decisions regarding the topic. A limitation in this research that may have skewed this data are irrelevant comments, such as "Where's the white guy?" This comment had nothing to do with climate change, and it would have been extremely difficult to filter this out of the sample of comments.

Table A

| | |
|---|---|
| Weighted Mean Sentiment ($\mu_{sentiment}$) | -0.118 |
| Weighted Standard Deviation ($\sigma_{sentiment}$) | 0.536 |
| Like Ratio (likes/dislikes) | 5.075 |

The second video, titled "Is It Too Late To Stop Climate Change? Well, it's Complicated." posted by Kurzgesagt [14], a German animation studio was selected for the same criteria that the previous video was selected, in that it was also a foreign source. The results are displayed in Table B. However, it was also selected because it was mainstream, as its YouTube channel has 14.6 million subscribers, and this specific video received 5.1 million views. The weighted mean sentiment was negative, but there was a large standard deviation, meaning the concluding factor was the like ratio. Since the like ratio was less than 10 (less than 10 likes for every dislike), there was sufficient data to correlate that with a foreign but mainstream source regarding climate change, there is likely to be a negative sentiment. However, it is important to note that even with a negative sentiment, this does not mean that the viewer base was against climate change. As aforementioned, it is possible that the viewers simply have a pessimistic view on how climate change will be solved, but these viewers may also support taking action on climate change. Therefore, a limitation to this research is not being able to pinpoint what a negative sentiment actually means in such a nuanced topic.

Table B

| | |
|---|---|
| Weighted Mean Sentiment ($\mu_{sentiment}$) | 0.247 |
| Weighted Standard Deviation ($\sigma_{sentiment}$) | 0.49 |
| Like Ratio (likes/dislikes) | 42.875 |

The third video, titled "Why People Don't Believe in Climate Science" posted by It's Okay To Be Smart, an educational science YouTube channel with 3.75 million subscribers sponsored by the Public Broadcasting Service (PBS) [15], was selected for being both mainstream — 1.67 million views — and an American media source. This case was an example of when the like ratio decided the sentiment, as both mean sentiment and standard deviation was neutral. Because the like ratio was greater than 100 (more than 100 likes per dislike), it was likely there was a positive sentiment, though it may be near neutral as the mean sentiment indicates. These results from the video indicate that an American, mainstream source may have a correlation with a more neutral or positive sentiment, as opposed to a negative sentiment.

Table C

| | |
|---|---|
| Weighted Mean Sentiment ($\mu_{sentiment}$) | 0.042 |
| Weighted Standard Deviation ($\sigma_{sentiment}$) | 0.199 |
| Like Ratio (likes/dislikes) | 135.08 |

## IV. Conclusion

Though this sentiment analysis algorithm seems to do exactly what it intends to do, it cannot differentiate between different negative or positive opinions regarding climate change without a human manually checking the results. However, because the scope of this research was small, the results of each video were verified manually. To improve upon the results of this research and also the algorithm, *Sentysis* is open-source, with a URL in A.1. Copies of the code can be made for free. The main improvement that could be implemented in *Sentysis* is machine learning, where the algorithm can be trained to easily identify a truly negative comment regarding climate change. Usually, sentiment analysis is combined with machine learning, as natural language processing (the processing that overarches sentiment analysis) deals with machine learning inherently. However, because that is such a large task, machine learning was not able to be implemented into this research.

## APPENDIX

A.1 *Sentysis* domain: Not available at this time. Either in development or hosted on a private domain. However, its open-source code can be accessed here: https://github.com/CodingPenguin/sentysis_flask and https://github.com/CodingPenguin/sentysis_angular.

A.2 Angular is a TypeScript framework for building single-page client applications. For more, visit https://angular.io/docs.

A.3 Flask is a Python microframework utilized for maintaining a web server. For more, visit https://flask.palletsprojects.com/en/1.1.x/foreword/#.

A.4 Youtube Data API v3 is a public API that allows developers to access the YouTube database for programs. For more, visit https://developers.google.com/youtube/v3.

A.5 Python is an object-oriented programming language that excels in sentiment analysis and machine learning. For more, visit https://docs.python.org/3/.

## ACKNOWLEDGMENT

Thank you to Vanielle Lee for guiding this research by providing your experience as a CTO and senior software engineer.

## REFERENCES

[1] M. Roser, H. Ritchie and E. Ortiz-Ospina, "Internet," *ourworldindata.org*, 2015. [Online]. Available: https://ourworldindata.org/internet#internet-access.

[2] SimilarWeb, "Youtube.com Analytics," *similarweb.com*. [Online]. Available: https://www.similarweb.com/website/youtube.com.

[3] N. Yiannakoulias, C. E. Slavik, M. Chase, "Expressions of pro- and anti-vaccine sentiment on YouTube," *Vaccine*, vol. 37, no. 15, p. 2057-2064, 2019. [Abstract]. Available: https://doi.org/10.1016/j.vaccine.2019.03.001. [Accessed Oct. 1, 2020].

[4] O. Uryupina, B. Plank, A. Severyn, A. Rotondi, A Moschitti, "SenTube: A Corpus for Sentiment Analysis on YouTube Social Media." Available: http://casa.disi.unitn.it/moschitti/since2013/2014_LREC_Uryupina_CorpusSentimentAnalysis.pdf. [Accessed Oct. 2, 2020].

[5] A. Balahur, "Sentiment Analysis in Social Media Texts," *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, p. 120–128, June 2013. Available: https://www.aclweb.org/anthology/W13-1617.pdf. [Accessed Sept. 29, 2020].

[6] "Descriptions of General Inquirer Categories and Use of Inquirer Dictionaries." Distributed by William James Hall at Harvard University. Available:

http://www.wjh.harvard.edu/~inquirer/homecat.htm. [Accessed Oct. 2, 2020].

[7] J. Bollen, H. Mao, X.J. Zeng, “Twitter mood predicts the stock market,” Oct. 14, 2010. [Online serial]. Available: https://arxiv.org/pdf/1010.3003&. [Accessed Oct. 2, 2020].

[8] A. Serbanoiu and T. Rebedea, “Relevance-Based Ranking of Video Comments on YouTube,” *Proceedings - 19th International Conference on Control Systems and Computer Science*, 2013. Available: 10.1109/CSCS.2013.87. [Accessed Nov. 30, 2020.].

[9] M. Rambocas, J. Gama, “Marketing Research: The Role of Sentiment Analysis,” *FEP Working Papers*, no. 489, April 2013. Available: http://wps.fep.up.pt/wps/wp489.pdf. [Accessed Nov. 30, 2020].

[10] N. Godbole, M. Srinivasaiah, S. Skiena, “Large-Scale Sentiment Analysis for News and Blogs.” Available: http://www.uvm.edu/pdodds/files/papers/others/2007/godbole2007a.pdf. [Accessed Nov. 30, 2020].

[11] C.J. Hutto, E. E. Gilbert. “VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text”. *Eighth International Conference on Weblogs and Social Media (ICWSM-14).* Ann Arbor, MI, June 2014. Available: https://github.com/cjhutto/vaderSentiment. [Accessed Feb. 2021].

[12] F.A. Nielsen. "A new ANEW: evaluation of a word list for sentiment analysis in microblogs". *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages, CEUR Workshop Proceedings*, vol. 718, p. 93-98, M. Rowe, M.Stankovic, A. Dadzie, M. Hardey. May 2011. Available: https://github.com/fnielsen/afinn. [Accessed Feb. 2021].

[13] CGTN America. "The Heat: Climate Change Challenges for 2021," YouTube, Dec 24, 2020 [Video file]. Available :https://www.youtube.com/watch?v=qF9

PpfBatBo&ab_channel=CGTNAmerica [Accessed: Feb. 2021].

[14] Kurzgesagt. "Is It Too Late To Stop Climate Change? Well, It’s Complicated," YouTube, Sep 29, 2020 [Video file]. Available: https://www.youtube.com/watch?v=wbR-5mHI6bo&ab_channel=Kurzgesagt%E2%80%93InaNutshell. [Accessed: Feb. 2021].

[15] It’s Okay To Be Smart. "Why People Don’t Believe in Climate Science," YouTube, Dec. 15, 2014 [Video file]. Available: https://www.youtube.com/watch?v=y2euBvdP28c&ab_channel=It%27sOkayToBeSmart. [Accessed: Feb. 2021].

[16] E. Boiy, M.F. Moens. “A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts”. August 13, 2008. Available: https://www.researchgate.net/publication/220479915_A_Machine_Learning_Approach_to_Sentiment_Analysis_in_Multilingual_Web_Texts. [Accessed Feb. 2021].