

Keywords on Online Video-ads Marketing Campaign: A Sentiment Analysis

Ronald R. Fernandez*
College of Commerce and Business Administration,
Research Center for Social Sciences and Education,
University of Santo Tomas, Philippines

Chin Uy
Graduate School, Research Center for Social Sciences and Education, University of
Santo Tomas, Philippines

— *Review of* —
**Integrative
Business &
Economics**
— *Research* —

ABSTRACT

The internet has made changes in the way people act that it even provides a forum for the expression of one's opinion. Consumers use social media to communicate and interact with other consumers. Due to the popularity of using social media in promoting products by marketers, opinion mining has become a trend as this examines the sentiments of consumers. Based on the ten most popular YouTube PH ads from January to June of 2017, the researcher identified the predominant keywords in the food and beverage industry video-ads extracted from the opinions of the viewers of these ads. The paper adapted the five stages in sentiment analysis (Rambocas & Gama, 2013) focusing on keyword extraction instead of sentiment detection. AYLIN text analysis using Rapidminer was utilized to classify the sentiments and determined the polarity and subjectivity confidence of the keywords. The study found that keywords to elicit emotion, words that refer to inspiring stories, and the integration of hit music trigger the curiosity of the viewers and make the video-ad popular. This finding may guide marketer in creating video-ads that will be popular and make a lasting impression on the viewers.

Keywords: sentiment analysis, keyword, online video-ad, marketing campaign

1. INTRODUCTION

The Internet has changed how people express their opinions. Since marketers need to monitor what people are saying about their brands, businesses need to follow suit. To cope with this ever-changing advancement, businesses utilize technologies like social media platforms so that they can better communicate and interact with their customers. These so-called opinions are important in performing sentiment analysis.

Sentiment analysis, also called opinion mining, is the field of study that examines the opinions, sentiments, emotions of people towards entities (Liu, 2010). These sentiments are expressed as either positive, negative, or neutral.

Nowadays, online platforms like YouTube are utilized by marketers to promote their products via video-ads and the like. Normally, these uploaded videos are not only viewed by people, but they also serve as an avenue to ferret out information through the comment section. Reading and retrieving these comments is quite easy if there were only a few of

them. However, if there were hundreds and even thousands, then the task becomes daunting. These sentiments from people are important to marketers to assess not only what people are saying about their brands, but it also serves as a gauge to examine the performance of their video ad placement as well. These comments provide a wealth of information for a variety of uses. One of the possible usages is the extraction of the predominant keywords from successful video-ads to determine the probable reason behind its success.

The objective of this paper is to determine the keywords that are found in the video-ads of the food and beverage industry in the Philippines through sentiment analysis to provide input to companies for product promotion. These identified keywords can also help marketers conceptualize better video-ads for a more successful marketing campaign for their products.

2. REVIEW OF RELATED LITERATURE

Video Ad comments

YouTube started in 2005, and since then, it has become the largest video-sharing website on the internet (Gill et al., 2007). It provides a facility for registered users to post comments in videos. This facility has become an avenue for communication related or unrelated to the video content (Madden, Ruthven, & McMenemy, 2013). Researchers are constantly studying YouTube to determine user behavior, measure video popularity for marketing purposes (Kousha, Thelwall, & Abdoli, 2012).

Video Ad Effectiveness

In advertising, it was found out that emotion plays a vital role in an advertisement (Kemp et al., 2012). Emotions trigger instant behavioral responses that are useful to marketers (Cohen, & Areni, 1991). And, emotions have an impact on satisfaction, costumers' loyalty, and decision making (Kwortnik, & Ross, 2007). Consumers spend more time viewing ads with positive content, for they create a positive feeling (Hirschman & Stern, 1999). Positive emotions improve the effectiveness of an ad. Although negative emotions also affect engagement to consumers, the level of influence is quite low (Kujur, & Singh, 2018). A strategy normally utilized in the effectiveness of the ad is the utilization of celebrity endorsers. Celebrity endorsement can be an effective tool to differentiate products in a market that is mature and saturated as long as the appropriate celebrity is located (Erdogan, 1999). These credible endorsers will make consumers attitude towards social media ads favorable (Samat, & Yusoff, 2014). Another factor to consider is a well-crafted story. According to Lundqvist et al. (2010), "a well-crafted firm story may create positive associations with a brand and ultimately increase consumers' willingness to pay for it." Music integration is another technique used in ads. Using hit music is an effective way to draw attention and induce positive replies from its listeners (Bruner, 1990).

Sentiment Analysis

Sentiment analysis, also called opinion mining, is the field of study that examines the opinions, sentiments, emotions of people towards entities (Liu, 2010). These sentiments are expressed as either positive, negative, or neutral. An example of positive sentiment is

“I love the film,” while “the movie is terrible” is an example of negative sentiment. A neutral sentiment expresses no feeling like “I watched the movie.”

Sentiment analysis, in its early stages, was used in studying the messages written in the stock boards to identify market sentiments (Das & Chen, 2001). The rise of machine learning and the availability of data sets to be trained were some of the factors that lead to its popularity (Pang & Lee, 2008). Its roots can be traced from different disciplines like psychology, sociology, and anthropology (Rambocas & Pacheco, 2018).

Sentiment Analysis in Marketing

In marketing, sentiment analysis has been utilized since time immemorial to capture the responses and feedback from their customers. Marketers have been studying the sentiments of their customer through surveys, comment cards, interviews, and focus groups (Rambocas & Pacheco, 2018). Although the sample size is small, it gets the work done. However, with the advent of the interactive environment of the internet, marketers can get the sentiments of their customers in large volume and in real-time. These interactive environments like social media have become an avenue for marketing, brand awareness, acquisition of new customers, and building of community networks (Blakeman & Brown, 2010). Even as early as 2007, the incorporation of academic libraries in YouTube was discussed (Webb, 2007). Though marketing was not mentioned, the usage of Youtube for marketing purpose was implied.

3. METHODS

3.1 Subjects of the study

The ten (10) most popular local ads shown in their platform for the first half of 2017 was released by YouTubePh (Tech and Lifestyle Journal, n.d.). The local ads were from various industries, namely: Telecommunications, Food and Beverage, and Personal Care Products. These large-scale business enterprises are in existence for the past 40 to 152 years. The video-ads in the study were extracted from this list. However, the study was limited to the video-ads that came from the Food and Beverage industry. The respondents of the study were people who posted in these YouTube sites.

3.2 Sentiment analysis mining stages

The study adapted the five stages in sentiment analysis proposed by Rambocas and Gama (2013) shown in Figure 1; namely data collection, text preparation, sentiment detection, sentiment classification, and presentation of output. Instead of sentiment detection, keyword extraction was the focus where the number of occurrences per words in the video-ads was determined.

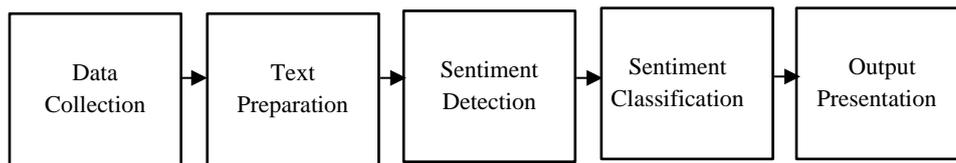


Figure 1. Sentiment analysis mining stages
Adapted from Rambocas, M., & Gama, J. (2013)

3.2.1 Data Collection

The comments on the food and beverage industry needed in the study were extracted individually and stored as excel files in CSV format. These pertained to comments made by viewers from January to June 2017.

3.2.2 Text Preparation

The individual CSV file was opened using a spreadsheet package. Initially, data extracted included information such as id, user, date, timestamp, likes, has replies, number of replies, replies id, replies user, replies date, and replies timestamp, comment text and replies comment text. The only information taken for the study were those on comment text and replies comment text.

The comments were merged to come up with a single comment column. These procedures were repeated for all the CSV files.

3.2.3 Sentiment Detection

In this stage, the sanitized CSV file was loaded to the Rapidminer Studio software for the mining of comments to come up with the predominant keywords and the number of occurrences.

The operators used are as follows:

- Read CSV: Reads data from a CSV file.
- Select Attribute: Selects a subset of attributes and removes the other attributes.
- Nominal to Text: Changes the type of selected nominal attributes to text.
- Trim: Removes leading and trailing spaces from the values of the selected attribute.
- Remove Duplicates: Removes duplicate text in the file.
- Tokenize: Splits the text in a document into a series of tokens.
- Transform Cases: Converts text in either lower or upper case.
- Filter stop words: Filters stopwords from the document based on a built-in stopword list.
- Generate n-grams: Creates terms represented by multiple words in a document.
- Filter tokens: Filter characters based on their length.
- Process Documents from data: Generate the number of occurrences from the generated text.

- Wordlist to data: Convert wordlist to the dataset in preparation for the excel output.
- Write Excel: Writes the data in an excel spreadsheet.

The excel output contains the keywords and the number of occurrences.

The entire procedure was repeated for all CSV files. After that, all excel file outputs were filtered manually for duplicates. The irrelevant keywords were purged, and then the top ten (10) keywords per ad were extracted.

The excel files of the individual ads that contain the top ten (10) keywords were then combined. To determine the top ten (10) keywords of the combined ads, the ranking was done via an excel pivot table.

The sanitized output now contains the top ten (10) keywords of the combined ads.

3.2.4 Sentiment Classification

In the sentiment classification stage, polarization is done to determine whether the keywords generated are positive, negative, or neutral. It will also generate the corresponding polarity and subjectivity confidence.

The stage starts with the reloading of the sanitized excel file generated from the previous stage. The Filipino words were changed to its English equivalent (such as *bata* to kid, *kilig* to elated, *ganda* to beautiful). The excel file was then loaded to Rapidminer studio with the integration of the AYLIEN text analysis.

The following are the operators used in this stage:

- Read Excel: Reads data from the excel file.
- Analyze Sentiment: AYLIEN text analysis operator. Extracts sentiments from the text that determines the insight of the author's emotion, whether the tone is positive, negative, or neutral. It also determines the text subjectivity (which means the author's opinion) or objectivity (which means the author is expressing a fact).
- Write Excel. Writes the data in an excel spreadsheet.

The excel output contains the ten (10) keywords with the sentiment analysis data.

3.2.5 Output Presentation

In the final stage, the graphs from Rapidminer Studio were copied, and all excel files were gathered into a new excel file in an organized manner for presentation purposes.

3.3 Ethical consideration

The study made use of data extracted from public YouTube sites. The comments were taken from viewers the identity of which is not known. Although the identity of the viewers who posted can be traced, the data underwent filtering processes such as deletion of id, user, date, and other information that may identify the viewers, leaving only the comment text and replies comment text columns. As such, the data used in the study were incognito.

4. FINDINGS

Extraction of keywords

There were only five companies from the food and beverage industry that were included in the recognition given by YouTube. These companies were coded A, B, C, D, and E. For each company; ten top keywords were extracted.

Table 1

Frequency of keywords extracted from individual video-ads

	Company A	f	Company B	f	Company C	f	Company D	f	Company E	f
1	Mclisse	148	buko	82	inspiring	29	song	166	lolo	575
2	Love	83	coconut	43	joyce	21	hahaha	63	nakakaiyak	364
3	commercial	75	bagong	17	tears	16	girl	40	love	132
4	ganda	62	bagong_biyak	17	good	12	love	36	nakakaawa	77
5	Cute	49	Artificial	14	god_bless	9	crush	25	gusto	62
6	hahaha	49	Available	14	story	9	ganda	23	gift	55
7	congrats	47	Fruit	14	thank	8	forever	14	miss	55
8	galing	43	Masarap	13	bata	7	kuya	13	masaya	52
9	bagay	33	ang_sarap	12	naman	6	ashley_ortega	11	bata	45
10	story	28	masustansya	12	aral	5	babae	11	cute	34

Table 1 shows that for the video-ads of company A, the word “Mclisse” which pertains to the artists in the ad got 148 mentions, while the word “story” got the least (28 mentions). For company B, the word “buko” which pertains to the product in the ad got 82 mentions while the word “masustansya” (nutritious) and “ang sarap” (delicious) both got 12.

The same table shows that for company C, the word “inspiring” got 29 mentions while the word “aral” (study) got 5. For company D, the word “song” got 166 mentions while the word “babae” which pertains to a woman got 11. For company E, the word “lolo” (grandfather) got 575 mentions while “cute” got 34.

The keywords from the five companies shown in Table 1 were combined into one excel file. The frequency of occurrence of the keywords in the file was taken collectively and shown in Table 2. They were ranked from 1 to 10 with 1 as the highest.

Table 2

Frequency and ranking of keywords extracted from the combined video-ads

Keyword	Frequency	Rank
love	5	1
Cute	4	3
Happy	4	3
Story	4	3

Cried	3	6
Good	3	6
Song	3	6
Bata	2	9
Kilig	2	9
Ganda	2	9

Table 2 shows that the word “love” got the highest rank. It appeared in the video-ads of all five companies included in the study. Tied for second place were the words “cute,” “happy,” and “story.” The three keywords that were in Filipino (bata, kilig, ganda) were included in 2 out of 5 of the companies’ video-ad.

Sentiment Analysis

Using Rapidminer studio, a sentiment analysis report was generated.

Table 3

Sentiment Analysis data result history

ExampleSet (10 examples, 4 special attributes, 1 regular attribute)

Row No.	polarity_confidence	subjectivity_confidence	polarity	subjectivity	Row Labels
1	0.804	0.880	positive	subjective	love
2	0.477	0.960	neutral	subjective	cute
3	0.857	0.935	positive	subjective	happy
4	0.541	1.000	neutral	subjective	story
5	0.490	0.994	neutral	subjective	cried
6	0.552	0.939	positive	subjective	good
7	0.456	0.994	neutral	subjective	song
8	0.537	0.992	neutral	subjective	kid
9	0.407	1.000	neutral	subjective	elated
10	0.635	0.705	positive	objective	beautiful

From among the keywords extracted from the individual and combined video-ads, the top ten keywords were generated. The top combined keywords are love, cute, happy, story, cried, good, song, kid, elated, and beautiful.

In Table 3, the word “happy” got the highest polarity confidence of .857, while “kilig” (elated) got the lowest (.407). Most of the keywords are subjective, as shown in the subjectivity column. Although there were more neutral than positive polarity, the table shows that there is no negative polarity. The word “story” and “kilig” (elated) got the highest subjectivity value of 1 while the word “ganda” (beautiful) got the lowest value of .705.

Table 4
Tabular presentation of statistics result history

Name	Type	Missing	Statistics		
Confidence_polarity polarity_confidence	Numeric	0	Min 0.407	Max 0.857	Average 0.576
Confidence_subjectivity subjectivity_confidence	Numeric	0	Min 0.705	Max 1.000	Average 0.940
Prediction_polarity polarity	Polynomial	0	Least positive (4)	Most neutral (6)	Values neutral (6), positive (4)
Prediction_subjectivity subjectivity	Polynomial	0	Least objective (1)	Most subjective (9)	Values subjective (9), objective (1)
Row Labels	Polynomial	0			Values beautiful (1), cried (1), cute (1), elated (1), ...[6 more] Details...

As shown in table 4, the average polarity confidence of the keywords is 0.576, while the subjectivity confidence is 0.940. It is also observed that there are more neutral than positive in the prediction polarity. However, subjective outperforms objective in the prediction subjectivity.

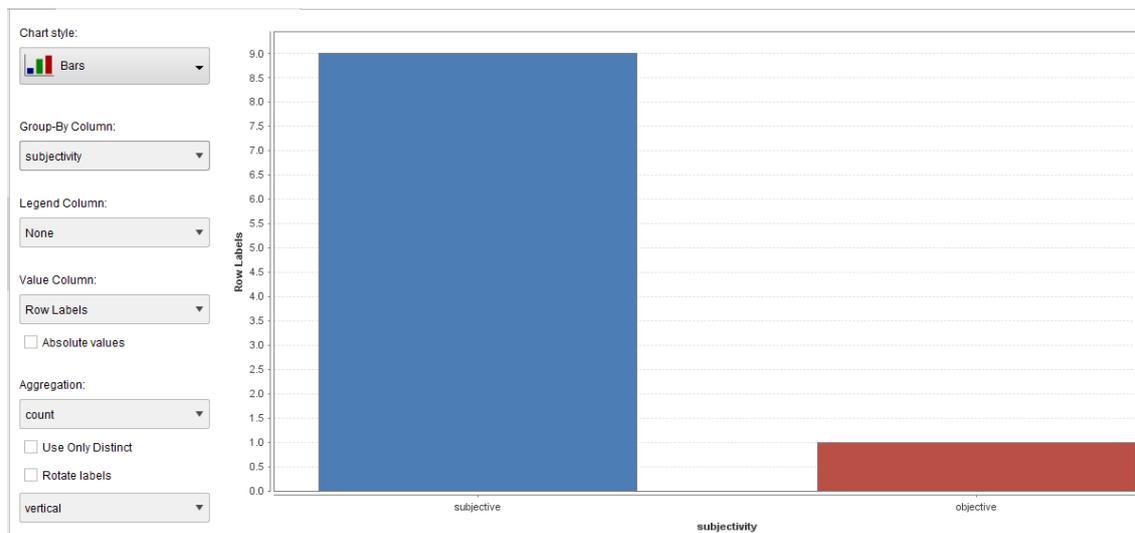


Figure 2. Graphical presentation of subjective versus objective comments

Figure 2 shows that the subjective value is 9, and the objective value is 1. That means, there is more subjective than objective comments.

5. DISCUSSION

This study shows the predominant keywords utilized in the ten (10) most popular YouTube ads in the Philippines from January to June 2017. Due to voluminous data available, the study concentrated on the Food and Beverage industry. There were five (5) large-scale business enterprises included in the study that has been in existence for the past 40 to 152 years.

Individual Video-ads

The predominant keywords per video-ad show that the word “Mclisse” which pertains to a celebrity love team in the country got the highest mention. This finding conforms with the study of Erdogan (1999) that celebrity endorsement can be an effective tool to differentiate products in a market as long as the appropriate celebrity is located.

The word “inspiring” which pertains to an inspiring story got the highest mention for the company’s video-ad. This result supports the findings that a well-crafted firm story may create positive associations with a brand (Lundqvist et al., 2010). In the same manner, the word “song” got the highest mention that pertains to the song used in the ad that piqued the interest of the viewers. The viewers were even asking for the title of the song. This result affirms the findings that using hit music is an effective way to draw attention and induce positive replies from its listeners (Bruner, 1990).

The word “lolo” or grandfather triggered the behavioral responses of the respondents because according to viewers, they missed their grandfathers. This finding is congruent to the study of Cohen and Areni (1991) that emotions trigger instant behavioral responses that are useful to marketers.

The use of the word “buko” or coconut that pertains to the product in the video-ad got the highest mention.

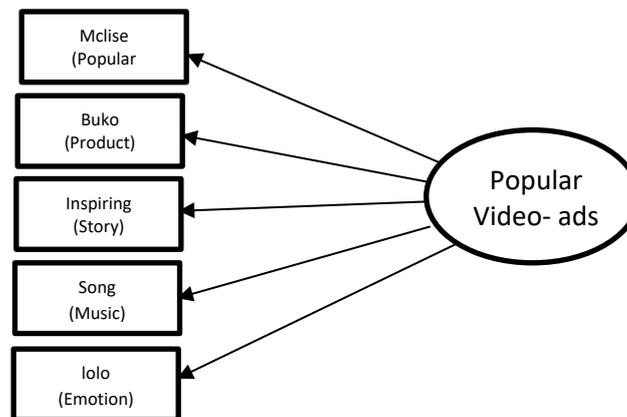


Figure 3 – Keywords of the Individual Video-ads

Combined Video-ads

When the top keywords were combined, the word “love” dominated as it was present in the video-ads of all the five companies. Other words that pertain to the emotions of the respondents such as “cute” and “happy” were present in four out of five video-ads. The word “cried” was present in three out of five. The word “bata” or kid inspired the viewers where some of them cried by the dedication of the kid in the ad. The word “kilig” elated the viewers by an exciting or romantic experience. This result affirms the findings of Cohen and Areni (1991) that emotions trigger instant behavioral responses that are useful to marketers. Even if others do not concur that the word “cute” is an emotion. However,

Sherman et al. (2011) stated that “the cuteness response is not yet well understood, but we believe it is an emotional response.”

The word “song” appeared sixth in the rankings that pertain to the music used in the ads that are parallel to the study of Bruner (1990) that states that using hit music is an effective way to draw attention and induce positive replies from its listeners.

Also, in the top ten keywords are the word “story,” “good,” and “ganda” or beautiful that pertains to the story of the ads.

Figure 4 shows the visual representation of the predominant keywords combined for all the ads and their respective associations in the study.

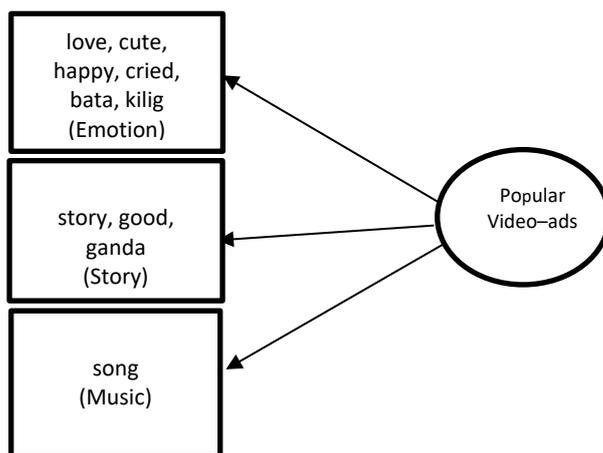


Figure 4 – Keywords of the Combined Video-ads

Data Result History, Statistics Result History, and Subjectivity Bar Chart

In the data result of the sentiment analysis using Rapidminer with the AYLEN operator, the result shows us that in the polarity (emotions of the viewers), there were no negative perceptions; that is, the viewers generally enjoyed the ads even though some of them were affected emotionally. There were more “subjective” than “objective” that tells us that the keywords given were the biased perceptions of the viewers. The polarity confidence is also quite high (.857), that confirms that the polarity prediction given by the system is accurate while the subjectivity confidence got the highest value (1) that reinforces the fact that the prediction given by the system is accurate and very subjective.

6. CONCLUSION

This study has contributed to the literature in two (2) ways. First, it enumerated the steps undertaken to extract comments from the YouTube website and the different stages and operators needed to come up with the keywords using Rapidminer, as well as the usage

of AYLIEN to come up with sentiment analysis. Researchers and practitioners can use this study as a guide in creating their text extractions and the development of their sentiment analysis.

Lastly, the study identified the keywords needed in producing popular video-ads. In the individual video-ads, keywords like Mclisse (celebrity), buko (product), inspiring (story), song (music), lolo (emotion) were the keywords that got the highest mentions. While in the combined video-ads, the keywords were love, cute, happy, cried, bata, kilig (emotion); story, good, ganda (story); song (music). From the keywords that were extracted, it is apparent that the use of celebrities, the integration of hit music, interesting/inspiring story, and touching emotions are factors that triggered the popularity of these video-ads. As a marketing practitioner, these keywords can be considered in making video-ads for marketing purposes, particularly for those that belong to the Food and Beverage industry.

For future research, it is interesting to know if the study will still be the same if more industries were included. And since the study covered only the video-ads from January to June 2017, it is also interesting to find out if the results will still be the same for July to December of the same year. An area that can also be considered is the comparative study of the keywords from the first versus the second half of the year to determine if there is a significant difference between the two.

REFERENCES

- [1] AYLIEN. (n.d.). In AYLIEN website. Retrieved November 10, 2017 from <https://aylien.com/>
- [2] Bruner, G. (1990). Music, Mood, and Marketing. *Journal of Marketing*, 54(4), 94-104.
- [3] Blakeman, K., & Brown, S. (2010). Social media: Essential for research, marketing and branding. *Bulletin of the American Society of Information Science and Technology*, 37(1), 47-50.
- [4] Cohen, J.B., & Areni, C.S. (1991). *Affect and consumer behavior*, in Robertson, T.S. and Kassarian, H.H. (Eds), *Handbook of Consumer Behavior*. Prentice-Hall, Englewood Cliffs, NJ, pp. 188-240.
- [5] Das, S., & Chen, M. (2001). Yahoo! for Amazon: Extracting market sentiment from stock message boards. In *Proceedings of the Asia Pacific Finance Association Annual Conference (APFA)*, 35, 43.
- [6] Erdogan, Z.B., (1999). Celebrity Endorsement: A Literature Review. *Journal of Marketing Management*, 15(4), 291-314.
- [7] Gill, P., Arlitt, M., Li, Z., & Mahanti, A. (2007). YouTube traffic characterization: A view from the edge. *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement—San Diego, CA*, pp. 15-28.
- [8] Hirschman, E.C., & Stern, B. (1999). *The roles of emotions in consumer research in Arnould, E.J. and Scott, L.M. (Eds). Advances in Consumer Research*, Association for Consumer Research, Montreal, pp. 4-11.
- [9] Kemp, E., Bui, M., & Chapa, S. (2012). The role of advertising in consumer emotion management. *International Journal of Advertising*, 31(2), 339-353.

- [10] Kousha, K., Thelwall, M., & Abdoli, M. (2012). The role of online videos in research communication: A content analysis of YouTube videos cited in academic publications. *Journal of the American Society for Information Science and Technology*, 63(9), 1710-1727.
- [11] Kujur, F., & Singh, S. (2018). Emotions as predictor for consumer engagement in YouTube advertisement. *Journal of Advances in Management Research*, 15(2), 184-197.
- [12] Kwortnik, R.J., & Ross, W.T. (2007). The role of positive emotions in experiential decisions. *International Journal of Research in Marketing*, 24(4), 324-335.
- [13] Liu, B. (2010). *Sentiment analysis and subjectivity*. In N. Indurkha & F. J. Damerau (Eds.), *Handbook of natural language processing* (pp. 627–666). Chapman and Hall: CRC Press.
- [14] Lundqvist, A., Liljander, V., Gummerus, J., & Riel, A. (2013). The impact of storytelling on the consumer brand experience: The case of a firm-originated story. *Journal of Brand Management*, 20(4), 283–297.
- [15] Luo, L., Wang, Y., & Han, L. (2013). Marketing via social media: A case study. *Library Hi Tech*, 31(3), 455-466.
- [16] Madden, A., Ruthven, I., & McMenemy, D. (2013). A classification scheme for content analyses of YouTube video comments. *Journal of documentation*, 69(5), 693-714.
- [17] Ortigosa, A., Martin, J., & Carro, R. (2013). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31, 527–541.
- [18] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1–135
- [19] Peslak, A. (2017). Sentiment analysis and opinion mining: Current state of the art and review of Google and Yahoo Search Engines' Privacy Policies. *Journal of Information Systems Applied Research*, 10(3), 38.
- [20] Rambocas, M., & Gama, J. (2013). Marketing research: The role of sentiment analysis (No. 489). Universidade do Porto, Faculdade de Economia do Porto. Retrieved from <http://www.fep.up.pt/investigacao/workingpapers/wp489.pdf>
- [21] Rapidminer. (n.d.). In Rapidminer website. Retrieved November 10, 2017 from <https://rapidminer.com/>
- [22] Samat, M.F., Hashim, H., & Yusoff, R.N.R. (2014). Endorser credibility and its influence on the attitude toward social media advertisement in Malaysia. *Review of Integrative Business & Economics*, 4(1), 144-159.
- [23] Sherman, G.D., & Haidt, J. (2011). Cuteness and disgust: The humanizing and dehumanizing effects of emotion. *Emotion Review*, 3(3), 245-251.
- [24] Tech and Lifestyle Journal. (n.d.). 10 most popular YouTube PH ads from January to June 2017. Retrieved 2017 from <http://techandlifestylejournal.com/10-youtube-ph-leaderboard-ads-january-june-2017/>.
- [25] Webb, P.L. (2007). YouTube and libraries: It could be a beautiful relationship. *College & Research Libraries News*, 68(6), 354-355.
- [26] Wu, X., & Bolivar A. (2008). Keyword extraction for contextual advertisement. WWW '08 Proceedings of the 17th international conference on World Wide Web, 1195-1196.
- [27] YouTube Comment Scrapper. (n.d.). In YouTube Comment Scrapper website. Retrieved December 8, 2017 from <http://ytcomments.klostermann.ca/>