NOVA IMS Information Management School

MGI

Master Degree Program in
Information Management

Factors driving engagement in social media – an analysis of the cognitive appraisal of video content.

Younes Abouljid

Dissertation

presented as partial requirement for obtaining the Master Degree Program in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

FACTORS DRIVING ENGAGEMENT IN SOCIAL MEDIA – AN ANALYSIS OF THE COGNITIVE APPRAISAL OF VIDEO CONTENT.

By

Younes Abouljid

Master Thesis presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Information Systems and Technologies Management

Supervisor: Dr. Nuno António

Co-Supervisors*:* Dr. Jurij Jaklič

November 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Younes Abouljid

Lisbon, 30 November 2023

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Dr. Nuno António from NOVA IMS and co-supervisor Dr. Jurij Jaklič from University of Ljubljana, School of Economics and Business for their suggestions and curation of the project idea and scope. Their help and guidance throughout the thesis conception and implementation, and for the previous semesters as a competent professor in their respective fields as I have learned a lot from their knowledge. In addition to all the professors whom I learned from and helped me develop both academically and personally.

I would also like to thank our research participants from various countries for sharing their opinions with us and taking the time to be involved in the research.

Finally, I humbly express my gratitude to my family, friends, and mentors that supported my journey in this European masters, helped me grow, discover my potentials, and contributed to what I've achieved so far. You are truly appreciated!

ABSTRACT

This research aims to elucidate the factors that drive engagement with video-centric social media content, focusing on the cognitive appraisal of such content. This study employs sentiment analysis, utilizing the analytical capabilities of Power BI to interpret data gathered through OpenCV machine learning models and user feedback surveys. The objective is to discern the emotional responses elicited by different types of engagement. The data collected from various sources was synthesized and standardized to uncover patterns, correlations, and key determinants of social media engagement and the emotions it triggers. By gaining insights into the emotional aspects triggered by media content, campaign managers can more effectively design and test their campaigns prior to deployment, thereby enhancing engagement and mitigating potential shortcomings.

KEYWORDS

Sentiment analysis, cognitive appraisal, social media, customer engagement, emotional response, video content

Sustainable Development Goals (SGD):

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

BI: Business Intelligence

EDA: Exploratory Data Analysis

ETL: Extract, Transform, Load

HCI: Human-Computer Interaction

ML: machine learning

NLP: natural language processing

SNS: social network sites

SVM: Support Vector Machine

TLM: thought listing method.

1. INTRODUCTION

Social media has become an essential part of modern communication and marketing (Schreiner, Fischer, & Riedl, 2019). With the increasing popularity of video content on social media platforms (Schreiner, Fischer, & Riedl, 2019), understanding the factors that drive user engagement and the success of marketing campaigns has become a critical issue for businesses and individuals alike. This study aims to investigate the specific aspect of social media engagement with a focus on short video format content.

The study of emotions, particularly emotions derived from facial expressions, started as early as the 1990s (Ekman, 1992). This practice has evolved from face-to-face experimentation sessions into an automated process done through modern technology. This process is enabled via facial recognition software coupled with emotion detection models. In this study, we will take advantage of advanced technological capabilities to uncover the dynamic relationship between emotional response and engaging with the new and heavily popularized short video content.

The objective of this research is to identify and understand the reasons why some video-centric posts receive a high level of engagement while others don't through an analysis of the dynamics of the interaction between user emotions, engagement, and content characteristics. In addition to creating an automated flow that enables the cognitive appraisal test of a designed campaign.

the research questions are:

1. What are the emotional factors that drive user engagement with video content on social media platforms?

2. Why some videos succeed and other fail in driving the engagement of social media users even if they are posted through the same entity?

Addressing these research questions will provide a study framework and a digitally maintained pipeline that aspires to drive insights from the dynamics of the interplay of emotional responses with engagement levels. Eventually, contributing to practical insights and further academic research in consumer psychology and marketing.

This research paper also aims to explore a research gap that exist in understanding the extent to which computational models accurately capture human emotions specifically in video content. To address this gap, this research paper aims to explore the alignment between the emotions identified through sentiment analysis techniques and those experienced by human participants.

One of the added values of this study is the focus on the new wave of short-form videos, along with the usage of recent technologies to achieve this analysis. This comparative study aims to uncover patterns and alignment between what the emotional models will predict and what actual participants feel about certain videos. The study will be performed on a single entity. This means that in a specific timeframe, we will analyze different videos of the entity and measure the level of engagement to understand why some videos succeed while others don't in engaging social media users, based on the emotions derived from each video.

2. LITERATURE REVIEW

Several studies have discussed the effect of social media content, particularly video formats, as a mediating effect that triggers an impact on viewers (Kulkarni, Karlo, Sharma, & Sharma, 2019). The behavioral engagement driven by the experience of interacting with a Social Network Site (SNS) content is what companies are investing in and studying to maximize their retention rate and attract new customers. Studies involving emotions are hard to measure due to their subjectivity and contextual landscape.

In this section, we will focus on breaking down concepts related to studying emotional response, user engagement, and their typology. Understanding all the different methodologies used in previous literature and their shortcomings, as well as giving context to the research questions and identifying the research gap from the previous work.

2.1. EMOTIONAL RESPONSE

In the context of studying sentiment analysis, the emotional response, also referred to as the cognitive response, is the impression the subject feels right after or during exposure to media content. In a study by Mauri et al. (2011), using various physiological measures such as "skin conductance, pupil dilation, blood volume pulse, respiratory activity, electromyogram, electroencephalogram" the viewers displayed signs similar to a flow state of solving a mathematical equation or dealing with a stressful problem that are characterized by high valence and strength (Schreiner, Fischer, & Riedl, 2019).

Emotions: Since emotions are a difficult concept to measure and understand, researchers use a popular model for emotional studies developed by Russel (2005) called the "Circumplex model of affect". This model measures emotions based on the valence or direction of the emotion, whether it is negative, positive, or neutral, and the intensity of the emotion or level of arousal. By understanding the emotion in a two-dimensional way, it could be manipulated by advertising companies and marketers (Kulkarni, Karlo, Sharma, & Sharma, 2019). For instance, the valence could be directed by choosing where to place an ad in between the show types (happy or sad TV shows). Also, the intensity of the emotion could be manipulated as well, showing that the most memorable ads are associated with high arousal levels.

Throughout the study of the literature about this topic, two main models are used by the industry's researchers to understand emotional response and its typology.

2.1.1. Russel's "Circumplex model of affect"

This model is a two-dimensional psychological framework that classifies and arranges emotional responses based on valence, or the direction of the emotion, and arousal, or the intensity of the emotion. Valence is classified into three categories: positive, negative, and neutral. The direction of the emotion reflects the emotional experience lived and the emotional response it generated. Emotions such as Rejoice and Happiness belong to the positive side of the spectrum, while emotions like anger and disgust fall on the negative side of the spectrum. The arousal refers to the tone of the emotion or the intensity of the emotional experience. Typically, it is scored based on a pre-defined range such as (-1 - +1) or (1 to 5). An example of an emotional response with a low arousal score would be calmness, and a high arousal emotional response would be anxiety. (Posner, Russel, & Peterson, 2005). Figure 2.1 depicts the arousal and valence described by Russel with a set of emotions that exemplify the four different parts of this model.

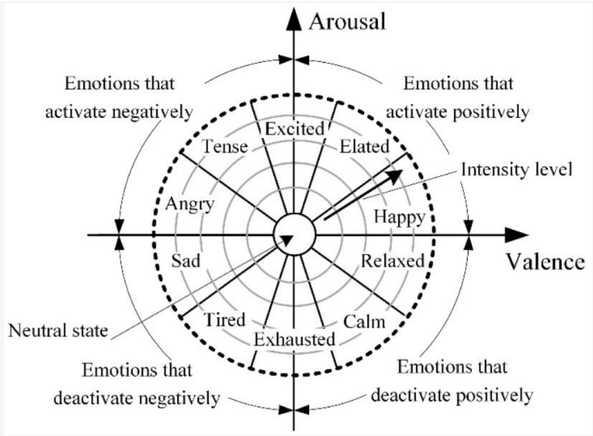


Figure 2.1 - Russel's circumplex model of emotions. (Russel, 2005)

2.1.2. Multimodal Sentiment Analysis

This model uses various modalities and artifacts to understand the emotional response in depth. Data is driven from images, texts, videos, sound, and physiological cues such as the heart rate and breathing rate. The multimodal uses more accurate emotions, typically typologized into six emotions: happy, angry, sad, surprise, disgust, and fear (Rao, Ahuja, Kansara, & Patel, 2021) as opposed to the traditional one developed by Russel where the valence is either positive, negative, or neutral. A very important characteristic of multimodal sentiment analysis is the ability to fuse, or what Bardzell et al. (2009) referred to as the triangulation method. This fusion allows for better accuracy of the emotional response studied. It takes results driven out of the text processed, the video frames analyzed, and the results of the heartbeat to see if they all indicate the same emotion in terms of typology, valence, and tone.

However, this ability for fusion comes with the problem of standardization and synchronization of the results across the different modalities. To overcome this challenge, Bardzell et al. (2009) attempted the standardization of the physiological modality by collecting the physiological rates, which were averaged and then standardized using the z-score to compute the population mean and standard deviation. In the controlled lab environment, a benchmark was set to a range of heartbeats per minute between 70 and 102. Thus, each physiological score (Heart Rate, Breath Rate) is a value representing the distance from the benchmark, whether it is higher or lower per video rating. The ANOVA analysis gives a high significance between the heart rate and the rating of whether the breath rate shows lower significance.

2.1.3. The six basic emotions

The choice of emotions to be analyzed in this study was found throughout most of the literature in the domains of sentiment analysis, affect computing, and human-computer interaction (HCI). Most of the studies were based on the work of Paul Ekman, a renowned psychologist and pioneer in discussing the relationship between emotions and the facial expressions represented by them. In his article Ekman argued that there are six basic emotions—later, a seventh emotion will be added in further literature—

that are universal and shared across all cultures through a consistent facial representation. These emotions are Enjoyment, Surprise, Fear, Disgust, Anger, and Sadness. His argument in the article was based on nine characteristics: (Ekman, 1992)

Distinctive universal signals: the author suggests that there are biological universal signals that humans perform using facial expression to express an emotion, even amongst remote societies that did not interact with the world civilization. For instance, a happy emotion is expressed through a smile, raised cheeks, and a reduction in the size of the eyes. These expressions were unique for each emotion and could be easily identified. The author also expressed that surprise and fear are the two emotions among the six that are hard to distinguish based solely on facial expression (Ekman, 1992).

Presence in other primates: this argument is based on the evolution theory, which states that these emotions not only exist in humans but could also exist in different primates and have a degree of expression through their physiological or behavioral aspects. Thus, besides their cross-cultural aspect, these six emotions are also cross-species (Ekman, 1992).

Distinctive physiology: these emotions have an effect on the physiological body and trigger a physiological response. For instance, the fluctuation of hormone levels and the change in heart and breath rate. These distinctive changes that occur due to the emotion felt cause a categorization among the emotions and reinforce the six universal emotions (Ekman, 1992).

Distinctive universals in antecedent events: there are universal events that trigger these emotions and thus their expression. For instance, injustice would trigger anger universally, and the loss of a loved one would cause sadness on the same level (Ekman, 1992).

Coherence among emotional responses: when an emotional response is expressed, it is not only done through facial expression but also through various expressive channels, such as the physiological one, that are unified and consistent. A fear emotion would cause a fight or flight response coupled with alertness, an increase in adrenaline, a heightened heart rate, and a short breath rate, as well as the accompanying facial expression (Ekman, 1992).

Quick onset: the activation of these emotions and their facial expression is done almost instantly. As the author exemplifies, a fight or flight response is triggered the minute fear is felt within. It is a quick response to the stimulus of the emotion (Ekman, 1992).

Brief duration: the lifespan of the emotion is dependent on the stimuli. It has a short duration, which comes as a response. It is usually diminished once the stimuli or the trigger are resolved, or adaptable if the trigger persists (Ekman, 1992).

Automatic appraisal: the reaction to the stimuli, internal or external, must be done at the highest level of speed, which uses primitive sensory events and subcortical neural circuits. (Ekman, 1992) “Emotional processing system… tend to use minimal stimulus representation possible to represent the activate emotional response control systems, which characteristically involve relatively hard-wired, species-typical behaviors and physiological reactions.” (Ekman, 1992)

Unbidden occurrence: the response or the appraisal resulted by the emotion is done automatically without a conscious effort behind it. It is an automatic response to the situation perceived; thus, it is characterized by an involuntary force of nature (Ekman, 1992).

2.2. Users Engagement

User engagement is the basis of a 300-billion-dollar industry known as content marketing (Schreiner, Fischer, & Riedl, 2019). The engagement could be described as the output of the emotional response after or during exposure to the media content. In the context of social media, engagement could be sharing the content (the most favorable form of engagement), liking, commenting, saving, and subscribing.

In an attempt to study engagement from media content, Schreiner et al. (2019) disregarded the physiological and psychological effects and focused more on the behavioral effects of individual viewers, meaning the action that the viewer takes after being exposed to the content, which can be objectively measured, and its positive effects are the desired outcome of marketing campaigns.

In another study by Nelson-Field et al. (2013), analyzing 800 Facebook ad videos using 16 different emotions, it was suggested that the high-arousal videos are shared twice as much as the low ones, and the positive-valence videos are shared 30% more than the negative-valence videos. However, engagement through commenting on posts doesn't seem to be affected by the arousal or the valence, according to another study by Yu in the F Journal on Media & Communication (2014).

The behavioral engagement can also depend on the media content characteristics. According to Schreiner et al. (2019), in their study, four content characteristics (topic, components, interactivity, and timing) have shown a degree of user behavioral engagement varying between low and high engagement.

The first content characteristic, which is the topic of the video, is also classified into several categories.

- Informative topics that provide information about the product, brand, or company.
- Transactional topics that include promotions, sales, discounts, and giveaways of certain products.
- Emotional topics that include a high level of romantic feelings as well as fear and humor.
- Transformational topics that include a message that appeals to the desires of viewers, mostly targeting their self-esteem.
- And lastly, entertainment topics, which are characterized by funny and humorous content.

The entertainment and emotional topics appeal to higher viewer engagement. As for the informative topics, they should be tuned down a notch to not bore the viewer. Also, transactional topics can elicit an emotional response of urgency due to the attractive deals presented (Schreiner, Fischer, & Riedl, 2019).

As for the second characteristic, the components, studies have shown that including the video component increased the rate of sharing and liking the video; however, there seems to be no effect on the commenting behavior. Whereas posts that include components such as mentions and hashtags increased viewer engagement slightly in the latter (Schreiner, Fischer, & Riedl, 2019).

The interactivity element of a post can also be a drive to engagement. A post that is in the form of a contest can increase engagement through likes and shares. On the other hand, a post that contains a question directed at viewers increases engagement through comments (Schreiner, Fischer, & Riedl, 2019).

The fourth content characteristic, timing, can also determine the level of engagement. The authors have concluded that posts made during the weekdays have a higher engagement level than those posted on the weekends (Schreiner, Fischer, & Riedl, 2019).

2.3. TYPOLOGY OF ENGAGED USERS

We understand now that behavioral engagement is the desired goal of every marketing campaign. It is a multibillion-dollar industry that prioritizes the feedback received from its users and consumers. In a study published by Kularni et al. (2019), the authors researched the different types of users who display various levels of engagement. Through the understanding of users, marketers could tailor their campaigns to target the seeders (highly engaged users) instead of a generic campaign that targets everyone and yet appeals to very few. Another interesting point that is mentioned in this study is the virality of marketing campaigns and advertising. According to the authors, there have been significant studies regarding the virality of advertising analyzed from a product's point of view (word of mouth), recipients (position in social networks, motivation, personality traits), and content (high level of entertainment, emotional experience). However, the authors expressed a research gap on the seeding strategies, or, in other words, the pre-selection of target consumers who are more likely to engage with the content and customize the message for them.

Analyzing cognitive responses is a measurement of the information processing abilities of an individual. "Consumers' processing of viral ads follows the popular 'Emotion-Cognition Model':" (Kulkarni, Karlo, Sharma, & Sharma, 2019), which means that consumers develop emotions from the ad, which in turn activate their cognitive appraisal, which incentivizes the act of sharing the ad. And this cognitive appraisal could be quantified and studied through users' comments, posts, and likes related to the ad. Throughout their research, the authors identified two types of cognitive appraisal: brand appraisal and ad content appraisal. There are also four factors that determine the success of a seeding strategy: content, network structure, behavioral incentives, and the seeding strategy itself. From their research, the findings were a typology of the users that interact with the media content and viral ad sharers. The 4 clusters are partitioned into 30.2%, 29.9%, 22.7%, and 17.2% of the total study pool. They were labeled as 'Active sharers', 'Brand-fanatic sharers', 'Content-hungry sharers' and 'Dormant sharers' together in a typology called ABCD, as shown in figure 2.2.

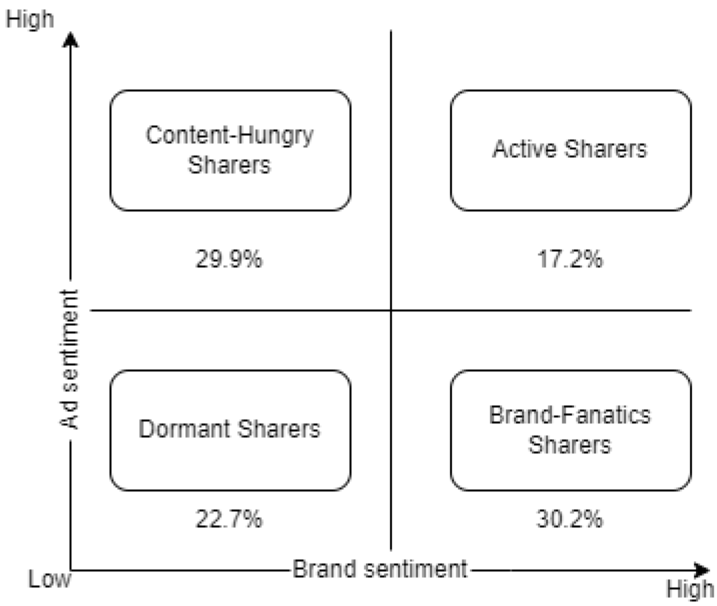


Figure 2.2 - Typology of viral ad sharers (Kulkarni, Karlo, Sharma, & Sharma, 2019)

Active sharers carry out the most sentiment for both the ad and the brand, while brand-fanatic sharers care more about the brand than the ad, and vice versa for the content hungry sharers. The dormant sharers simply do not care about the brand or the ad (Kulkarni, Karlo, Sharma, & Sharma, 2019).

The active sharers, or the most favorable cluster, which best seeds the viral campaign, are hard to find and relatively low compared to the rest of the users. In this study, 17.2% of the population was affected; however, the choice of these seed consumers who will potentially drive the virality of the ad campaign has to be done with precision. According to the classical Pareto principle (the law of the vital few), 20% of users are expected to carry 80% of the load to propagate the message. Sometimes companies epidemic viral campaigns tend to fail due to the broadness of the spread of the message and a general lack of proper segmentation (Kulkarni, Karlo, Sharma, & Sharma, 2019).

Predictors of virality are the importance of “message-content” particularly the emotional content, psychological (the need to belong), and tie strength as a feature of SNS.

Also, existing literature suggests that the incorporation of an embedded brand may distract the consumer from the message of the ad. Too much information regarding the brand embedded in an ad may lead to a negative outcome from sharing the viral ad; most consumers tend to focus on the content of the ad.

Newer literature suggests the opposite. A strong brand sentiment enhances the reception of the ad message, therefore increasing the sharing probability. It is also supported by the “classical ‘Reciprocal Mediation’ model of advertising effectiveness proposed by MacKenzie et al. (1986), which hypothesizes a reciprocal relationship between a consumer's ‘ad’ and ‘brand’ information processing behavior.”

The insights from this study help brand managers, advertisers, and marketers understand their user-generated content and devise effective strategies to meet the profile of the seeders in order to design a successful viral ad campaign.

2.4. OPINION MINING

Sentiment analysis, or opinion mining, started as early as the 1950s in the analysis of written documents (Rao, Ahuja, Kansara, & Patel, 2021). In his book Bing Liu described sentiment analysis as the “field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes” (Liu, 2012).

But since then, sentiment analysis has developed to analyze different artifacts using new techniques such as Natural Language Processing (NLP) and Machine Learning (ML), which have increased performance and accuracy. Organizations use sentiment analysis to gain competitive advantage, increase sales, and develop strategies that are more tailored and focused based on the opinions of their customers (Rao, Ahuja, Kansara, & Patel, 2021). There are two types of sentiment analysis techniques. Feature-based technique that extracts a special feature from a textual document. Document-based technique that derives the overall polarities of a text document (Kulkarni, Karlo, Sharma, & Sharma, 2019).

Initially, the traditional way to analyze the cognitive response is the ‘thought-listing method’ (TLM). While the modern way to analyze the cognitive response is through sentiment analysis driven by psycholinguistics and NLP (Kulkarni, Karlo, Sharma, & Sharma, 2019).

The TLM method is the traditional way to measure emotional response to video content or social media content in general. It is ad hoc based and has a subjective nature due to its indicators and intuition. The way it works is to write immediate thoughts during or after exposure to media content. Then, the experimenter classifies these thoughts based on pre-set criteria and sees which category they fit the most. Then, they proceed to calculate the score of each category, either by a simple addition or, if there are levels of importance, they use a model that weights each response to an indicator and eventually sums up and averages the scores to predict the attitude. Typically, the TLM has been used to compute a “valence-weighted aggregate index of cognitive responses." The shortcoming of this method is that it has measurement problems. Mainly, it captures the direction of the attitude, whether positive or negative, but fails to capture the intensity of the cognitive thought (Kulkarni, Karlo, Sharma, & Sharma, 2019).

Following the literature review exploring emotions, emotional response, user engagement, and opinion mining and their interrelation, this research paper aims to investigate whether the emotions detected via sentiment analysis methodologies align with those perceived by human observers. The methodology proposed will test the following hypotheses: H0 suggests that there is no significant difference between the emotions identified through sentiment analysis and those experienced by human observers, while H1 proposes the opposite. Throughout the research paper methodology and results discussion, we will try to validate the efficacy and accuracy of sentiment analysis techniques in capturing the nuances of human emotions. Through this investigation, we aim to contribute to a deeper understanding of emotional responses to video media content and their implications for user engagement in digital environments.

3. METHODOLOGY

This research design was designed to determine the reason behind the engagement success of social media videos. The focus is on the short-form video content that got heavily popularized after the 2020 lockdown. Amongst the top three leading platforms in this style of video content (Instagram reels, YouTube shorts, and Tik Toks), Tik Tok remains the leader in terms of downloads, usage, and engagement. According to Sensortower, a market-Leading Digital Intelligence in the first quartile of the year 2023, this platform has recorded the highest user engagement, with over 7 hours/week and 120 sessions/week per user, as shown in figure 3.1. (Abraham Yousef, 2023)

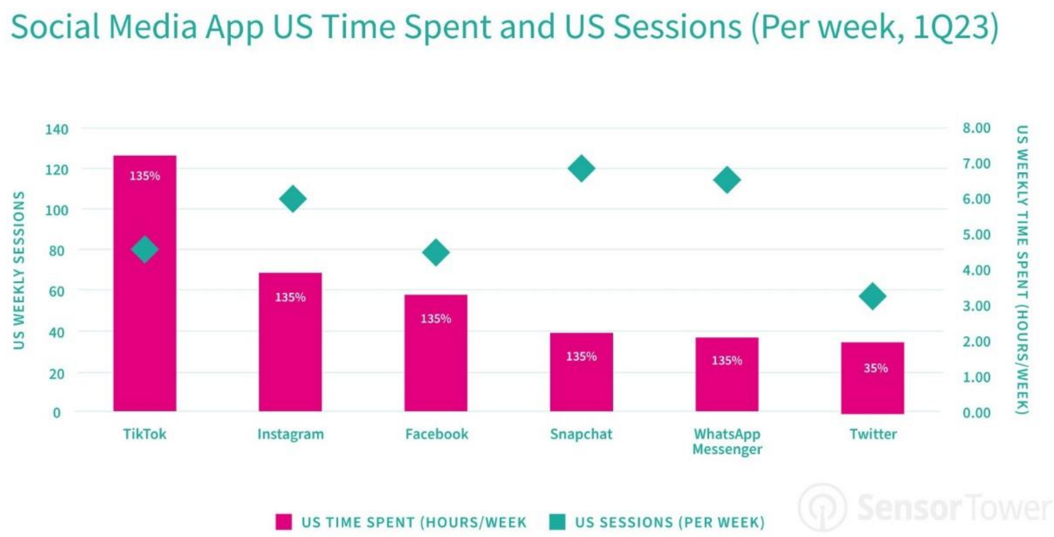


Figure 3.1 - Social Media engagement by 1st quartile of 2023 (SensorTower, 2023)

The short-form video content formula is characterized by a short duration, typically between 15 seconds and 1 minute, with heavy sound and video effects, bold texts, and creative filters, as well as creating a collaborative ecosystem between its users. (Singh, 2023)

This style of content got heavily popularized with the current fast pace of life that limits the user's attention span to consume short and digestible content that offers a variety of topics in a fun and expressive way (Singh, 2023). In order to understand the reason behind the success of this formula as well as how to survive in this current landscape, this study focuses on the analysis of the cognitive appraisal that is generated from this content.

The diagram shown in Figure 3.2 describes the research design process used to perform a comparative analysis between the machine learning models that analyze the videos selected through the TikTok platform and a survey-generated dataset that humanly analyzes the video by focusing on the emotional response. A python script is developed in Jupyter notebook to run the machine learning models against the videos selected for analysis. The results are exported into a CSV file which will be uploaded into PowerQuery a built-in ETL whitin the PowerBI suite. The data models consist of survey generated responses and machine learning models predicted emotional results. Then, derive the factors from the visual dashboards that drive the engagements, test the accuracy of the cognitive appraisal, and infer content characteristics factors of the brand-new style of videos dominating the social network sites.

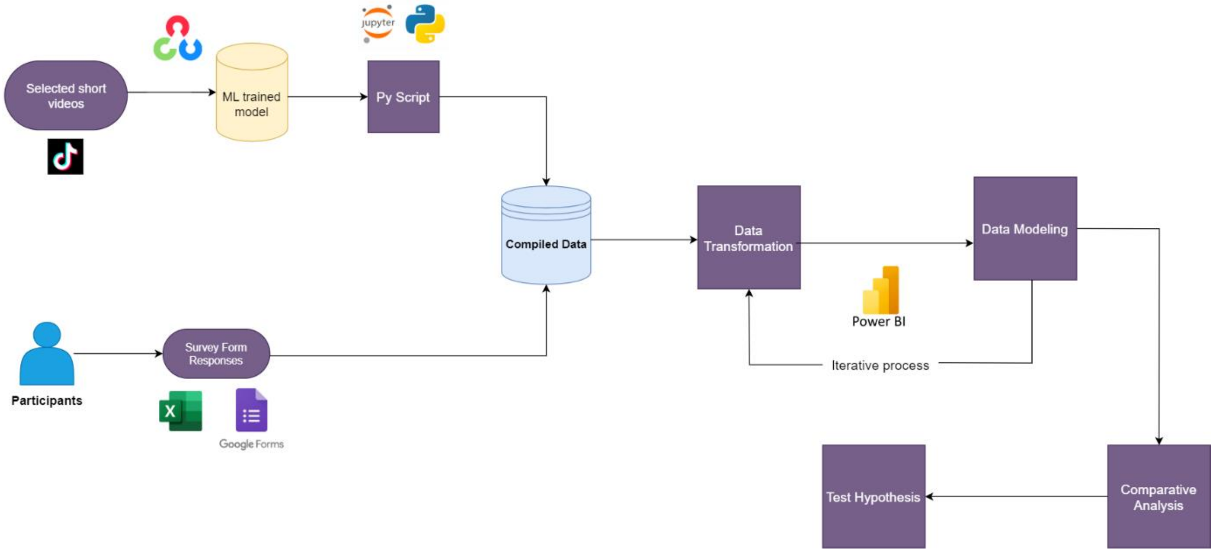


Figure 3.2 -Diagram of the research design

3.1. DATASET

Part of this study is to focus on a single entity. In other words, the choice of a single profile or influencer and seeing the fluctuation in engagement of their content over a specific period of time. Then, through sentiment analysis of the cognitive appraisal, infer the factors that resulted in the difference in the level of engagements.

The choice made in this study was for an influencer who is the holder of the Guinness world record for the most viewed video on the Tik Tok platform, reaching 2.2 billion views as of March 15, 2022 (Rodriguez, 2022). The profile goes by the name of Zach King, as seen in figure 3.3.

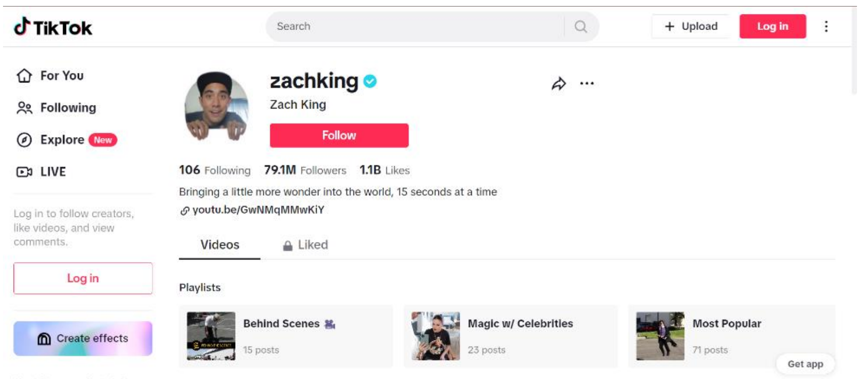


Figure 3.3 - Zach King TikTok profile (Tiktok, 2023)

The data was extracted during the month of April 2023, in the first month of this study. For that time period, 6 videos were extracted with different levels of engagement (see table 3.1), which would be the basis of this analysis both through the machine learning models developed by Octavio et al. (2019) and the survey designed and shared across different participants from different countries.

Table 3.1 - Videos selected for analysis from Zach King profile across the month of April.

Zach King Videos for the month of April 2023 Statistics	
Videos (ordered by chronological order)	Engagements statistics
(no title)	date: 05/04/2023 views: 25 800 000 likes: 1 300 000 comments: 3 372 saved: 38 700
When the day needs me anywhere and everywhere, #CeraVe SPF is what I wear so I can #FaceItLikeADerm 🥼 🌀 #CeraVePartner	date: 07/04/2023 views: 97 100 000 likes: 907 500 comments: 2 627 saved: 14 100
This is Puzzling 🧩	date: 14/04/2023 views: 135 600 000 likes: 8 200 000 comments: 17 500 saved: 366 800
Can someone please explain how the mirror knows there is an egg	date: 19/04/2023 views: 11 200 000 likes: 696 200 comments: 3 899 saved: 23 300
Life isnt always as it seems @sofyank96	date: 27/04/2023

	views: 80 900 000 likes: 5 300 000 comments: 16 800 saved: 164 600
spent the last 6 months making this, hope you enjoy the trailer. Watch the entire short film in my profile	date: 28/04/2023 views: 2 400 000 likes: 181 800 comments: 1 497 saved: 11 500

This dataset of videos also expressed a variety of content characteristics, which we discussed in the previous section, such as involving celebrities and other influencers, brand advertisements, and visual and sound effects. In addition to different topics that vary from promotional, sales, informative, and entertainment.

3.2. MACHINE LEARNING MODELS

3.2.1. OpenCV

The choice to use OpenCV technology as the main ML model to extract emotions based on facial expressions from videos came after several iterations and technological tools tryouts faced with dead ends such as experimenting with Microsoft Azure Cognitive Services. OpenCV refers to the Open Source Computer Vision Library. It is an open-source project licensed under the Apache 2 licensed products which makes it available for use for individual and commercial purposes (OpenCV team, 2023). The application of OpenCV algorithms as described in their official documentation on the GitHub platform could be used for the following: “to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc” (OpenCV team, 2023)

For this study, the model chosen was developed by b-it Bonn-Aachen International Center for Information Technology and Hochschule Bonn-Rhein-Sieg under the research paper published “Real-time Convolutional Neural Networks for Emotion and Gender Classification” The reported accuracy of this model was 70% for the emotional dataset and 96% for the gender dataset, with the goal of a real-time interactivity that matches human performance in analyzing emotions based on facial expressions by creating the best accuracy over a number of parameters (Octavio, Ploger, & Matias, 2019).

Regarding the emotional dataset, the developers used the Paul Ekman classification of the six basic universal emotions, which are Happy, Sad, Angry, Disgust, Surprise, Fear and the seventh added was a Neutral state of emotion.

However, there are some shortcomings described in the research paper describing the model. A major shortcoming is related to the accuracy of the model since, according to the authors, through the visualization of the confusion matrix, there is a faulty interpretation of the angry emotion with the disgust emotion, as well as the occurrence of confusion between the fear and sad emotions. This confusion typically happens when the subject is wearing glasses more frequently, especially dark glasses, since the model classifies it as a frown, which belongs to the angry emotion. Another shortcoming expressed by the authors is the bias of the model since it was trained using mostly western populations, which could reflect a classification that is rather inaccurate using eastern populations with eastern culture characteristics. (Octavio, Ploger, & Matias, 2019)

The models used follow a Haar classifier algorithm. An Haar classifier "is a machine learning object detection program that identifies objects in an image and video" (Mittal, 2020). It was first introduced by Paul Viola and Michael Jones in their 2001 paper, "Rapid Object Detection using a Boosted Cascade of Simple Features" (OpenCV, 2023). The algorithm needs both positive images, which are images with faces, and negative images without faces to train the classifier. The way the Haar classifier works is described in figure 3.4, where it moves through an object detector in a series of stages where, at each stage, the classifier decides if a region of pixels is positive or negative. It then discards the negative regions in which there is nothing of interest and moves on to the next stage. At the end, along with the model, it creates a system vision object able to analyze the image frames.

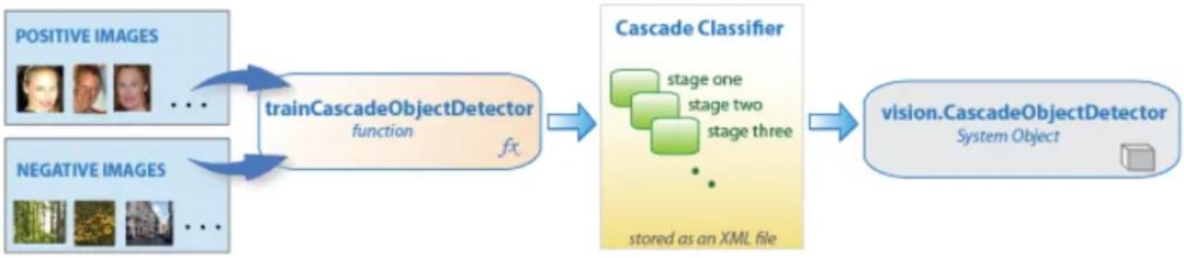


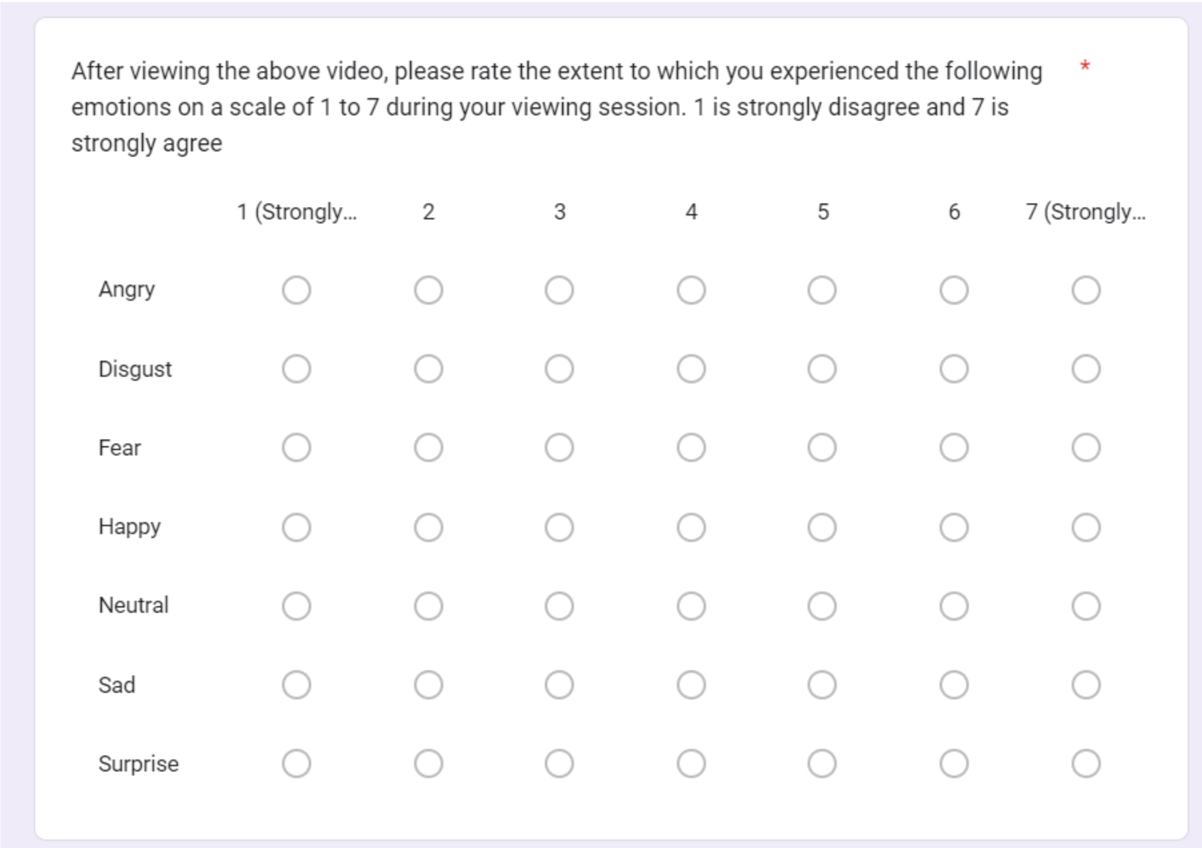
Figure 3.4 - flowchart describes the Haar cascade classifier design (Mittal, 2020).

3.2.2. Survey

The survey was generated through the Google Forms application and then manipulated for an initial phase through Microsoft Excel to clean up data related to scoring the emotions. Thus, simplifying the results and preparing them for import in the Jupyter notebook and Power BI for further analysis.

This survey was conducted via convenience sampling in which 1274 participants responded to the survey from different countries and various age demographics. The order of the videos was set by the researcher and the responses were all treated in a synthesized manner.

The purpose of the survey is to serve as a comparative element in the study between the scoring of the ML model of the emotions detected through video frames and the human emotions that are collected after viewing the videos through the question that uses a Likert scale from 1 to 7 to record the arousal and intensity of the emotion felt by the viewer. This question is displayed in figure 3.5.


After viewing the above video, please rate the extent to which you experienced the following emotions on a scale of 1 to 7 during your viewing session. 1 is strongly disagree and 7 is strongly agree *

	1 (Strongly...	2	3	4	5	6	7 (Strongly...
Angry							
Disgust							
Fear							
Happy							
Neutral							
Sad							
Surprise							

Figure 3.5 - the survey grid to collect emotions from videos watched.

The results collected through the survey were manipulated in an Excel file in order to perform the first steps of cleaning the data before importing it into the Power BI platform and performing further exploratory data analysis techniques and inferring results from it.

In addition to the scoring recorded by the Likert scale, other questions were asked in order to get a comprehensive understanding of the state of emotions experienced, its context, the reasons behind it, and the relationship of the viewer with the use of social media video content.

The survey template is attached in the Appendix A section for reference.

3.3. EMPIRICAL STUDY

The first part of this study was to analyze the emotional response through an ML model and extract emotions in terms of the frame count that exists in each video. To achieve these results, we created a separate virtual environment that includes relevant packages and libraries such as CV2, Keras, Pandas, Numpy, and Matplotlib.

In the notebook of analysis, we first loaded the emotional recognition model and the face detection model. To analyze the videos mentioned in the dataset, a constructed loop block breaks down the video into frames and processes each one using pre-trained face detection to extract the facial features, then predicts the emotion using an emotion recognition model, and then stores the emotion with the associated frame in a Pandas data frame, which will be further manipulated for analysis. Several steps were taken to achieve this result:

1. The model converts the frame to grayscale for faster processing.
2. It uses then the Haar cascade classifier to detect faces in the frame.
3. A loop will run through the faces that exist in each frame and predict the emotions out of them.
4. It adds the emotional analysis to the data frame.
5. After it runs through each frame, the video file is released, and all windows are destroyed.

So now we have a data frame that has all the frames read from the video with an emotion associated with them. However, there is a huge redundancy problem that needs to be solved since multiple similar emotions were detected across different video frames. To overcome this, we created another data frame that grouped the frames by the emotions and added them in a column called frame counts. Later, the column frame was dropped from the original data frame, and we merged the two of them into the emotion's column. Finally, we reset the index to make 'Emotion' a column again, which gives a final data frame as shown in Table 3.2, ready to be plotted and exported to be used in Power BI for further analysis. This process was repeated for all six videos.

Table 3.2 - Example of a final state of processed data frame.

Emotions	Frame Count
Angry	66
Disgust	1
Fear	17
Happy	205
Neutral	100
Sad	46
Surprise	2

To better visualize the results, we used the Matplotlib library as expressed in figure 3.6 to see the presence of emotions in relation to the frame counts. This plot helps as well in comparing the findings with the survey results, which we will discuss later.

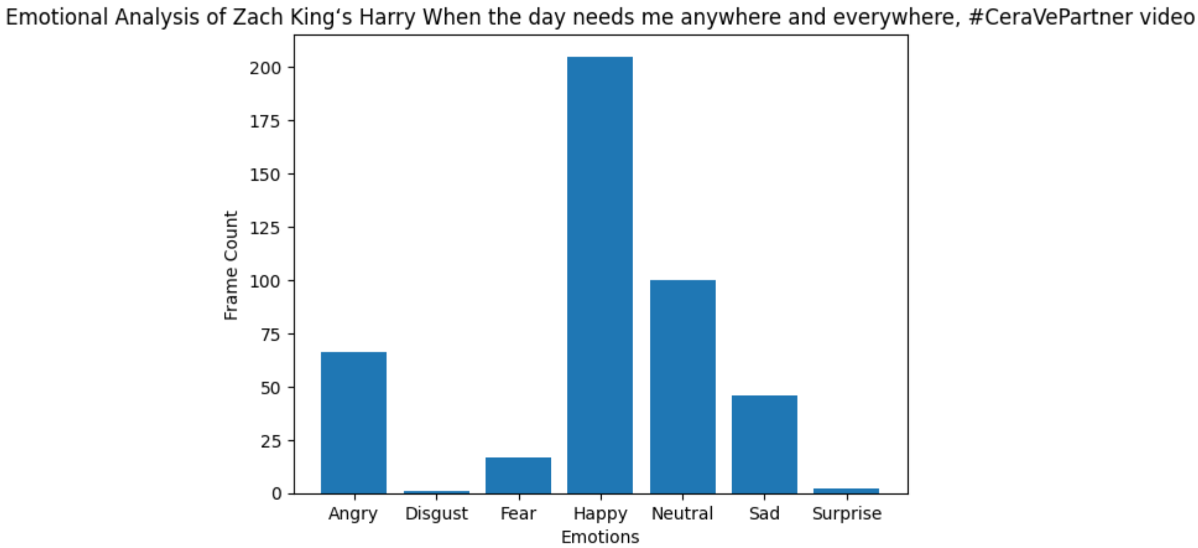


Figure 3.6 - plot of the emotional analysis data frame.

The same analysis process was iterated through all six videos, resulting in varying results on emotion recognition. The results are plotted in figure 3.7.

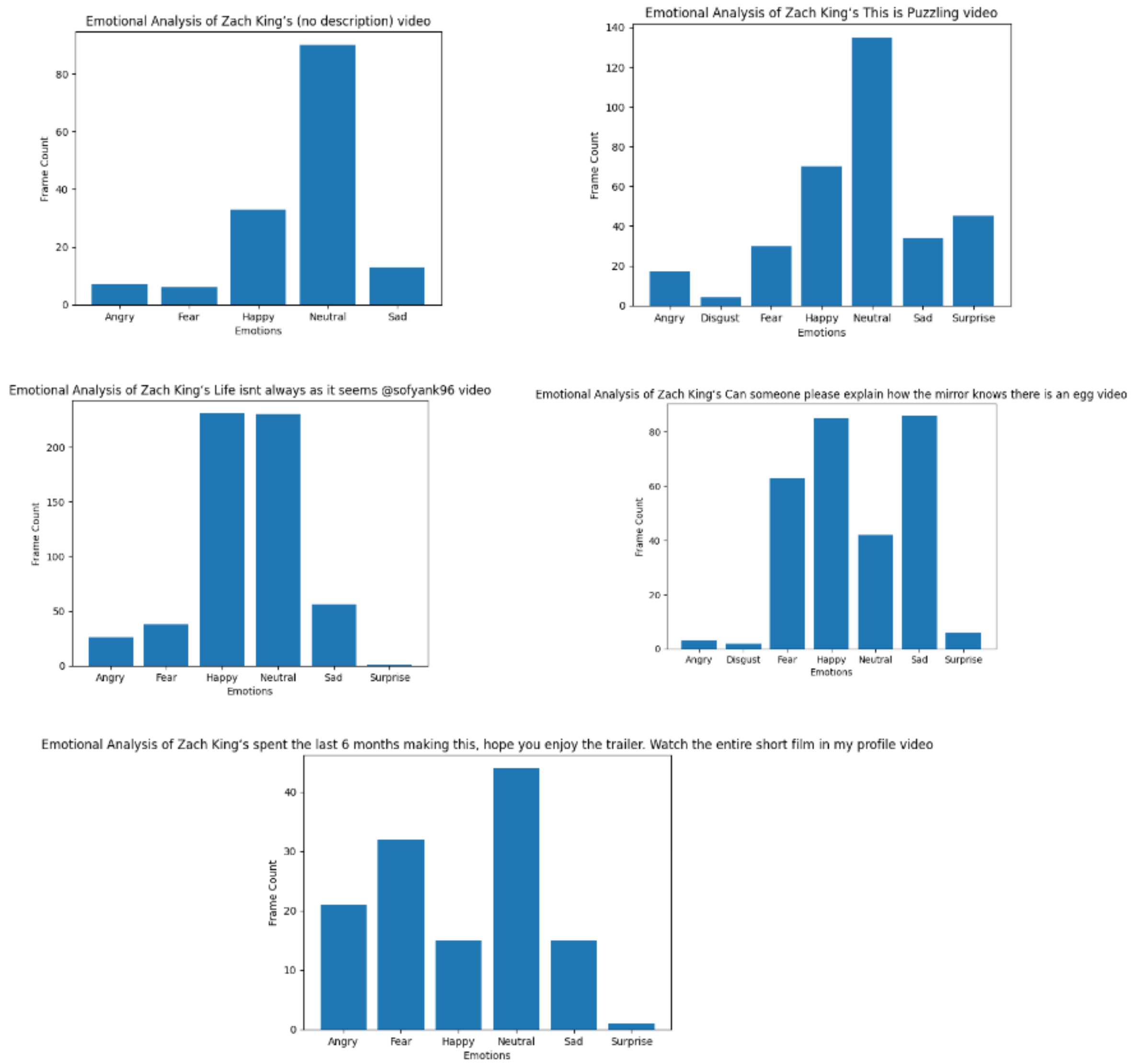


Figure 3.7 - Plots of the emotional recognition calculated by OpenCV model of the selected videos.

Notice that in some plots, we do not see all seven emotions extracted. This either goes back to the absence of that emotion in the video analyzed or the accuracy rate of the model and its shortcomings that we discussed in the previous section and how it can sometimes confuse some emotions, such as anger and disgust.

After plotting and visualizing the modeled data frames, which represent the emotions predicted by OpenCV, we exported them to six respective CSV files, which will then be imported to the Power BI platform for the comparative study.

Explanation of the survey study and integration of both findings from the survey and the ML model into Power BI:

ETL Power Query: The survey data was exported from the Google Forms application to a spreadsheet, which was then imported into Power Query. Power Query is an ETL (extract, transform, load) tool integrated in the Power BI ecosystem, and it allows the transformation of data to better visualize it.

The first step in this process is to import all the csv files generated by the ML models and the survey, plus an additional file in which we put the statistics of engagement extracted straight from the creator page.

Then, through power query, we append all the model's generated csv into a single table; each one has an identifier, and a frame count corresponds to the emotion predicted.

The second step was to manipulate the survey-generated data. In addition to the Likert scale scoring emotional response, other questions were present in the survey to understand the behavioral motives and the context behind the emotions captured. Examples of these questions were demographics, such as the age and country of the participants. As well as the reasons and time spent on social networking sites, the choice of content, and its characteristics. To clean up the noisy data, we change the type of some columns, such as changing from strings to numbers, in case the column has a numerical inference. Split the long spreadsheet into six tables that refer to the six videos. Each table contains nine columns that span the emotional response score and the first interaction recorded. Then, transpose the emotion columns into rows to match the model table format. Then, in each column, we added the id of the video recorded to prepare for the star schema model represented in figure 3.8.

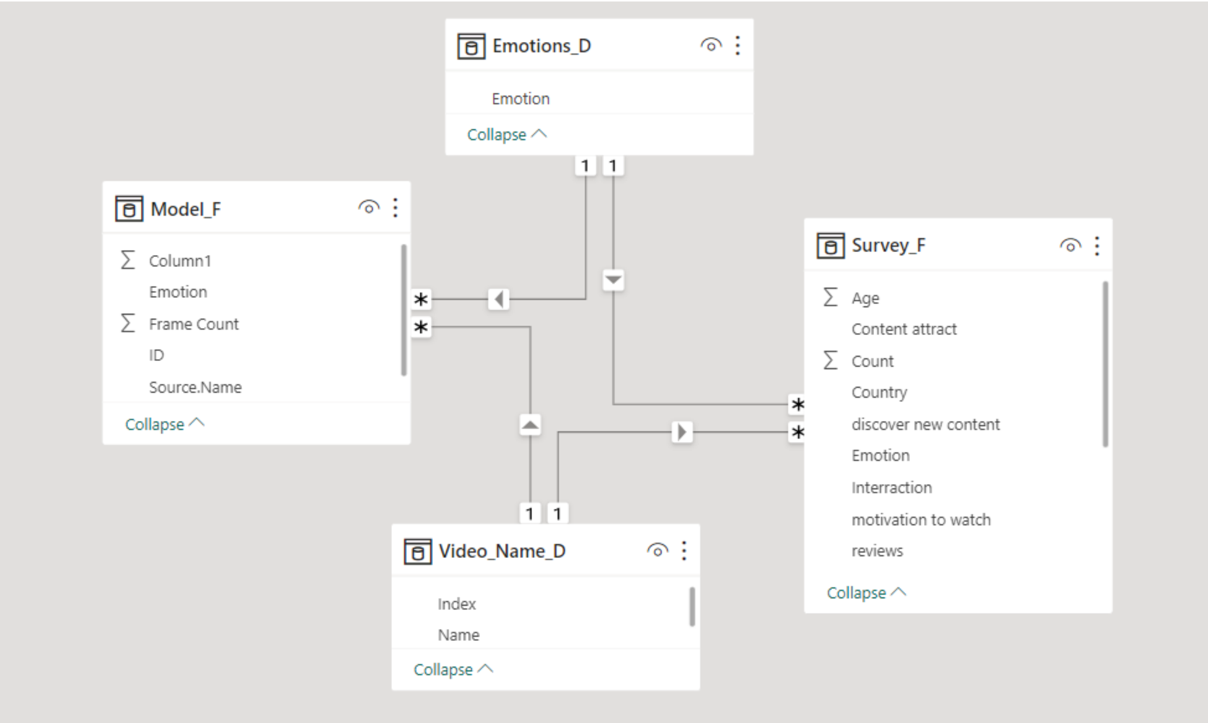


Figure 3.8 – Data model

After cleaning the data, we normalized the two fact tables: one for the survey and one for the model, using the dimensions of video name and emotion.

Normalizing values:

The ML model fact table has emotion values represented by frame count that ranges from 0 to N, and the survey fact table has emotion values represented by a Likert score that ranges from 1 to 7. Since both fact tables have different scales, a normalization process is mandatory to compare both fact tables with different scales.

To normalize the video scores, we implemented the min-max formula, represented as: $Xscaled = \frac{X - Xmin}{Xmax - Xmin}$. It scales the numerical values to a range between 0 and 1. It keeps the data proportional and preserves the relative order of data points.

In Power Query, we added two measures to each fact table that calculate the minimum value of the count column for each row, and we used the ALLEXCEPT function to remove all filters from the table except for the 'Video name' column, which ensures that the calculation considers the 'Video name' as a grouping criterion. The same logic applies to calculating the maximum value of the table. Then the DIVIDE function takes a detailed value, which is our x, and assigns the maximum value found in the 'Count' column per row minus the minimum value over a subtraction of the maximum value and the minimum value.

4. RESULTS AND DISCUSSION

The study findings were drawn from analyzing the reporting dashboard created to answer the research questions.

First, a comparative analysis of the emotional response between the ML model and survey responses. In this section, we will plot both normalized scores from the ML model and the survey into a clustered column chart where the x axis are the emotions, the y axis are the scores, and the legend indicates the orange color for the model values while the blue color is for the survey values. We will consider the top 3 emotions from each video and see their degree of alignment as displayed in figures 4.1 and 4.2.

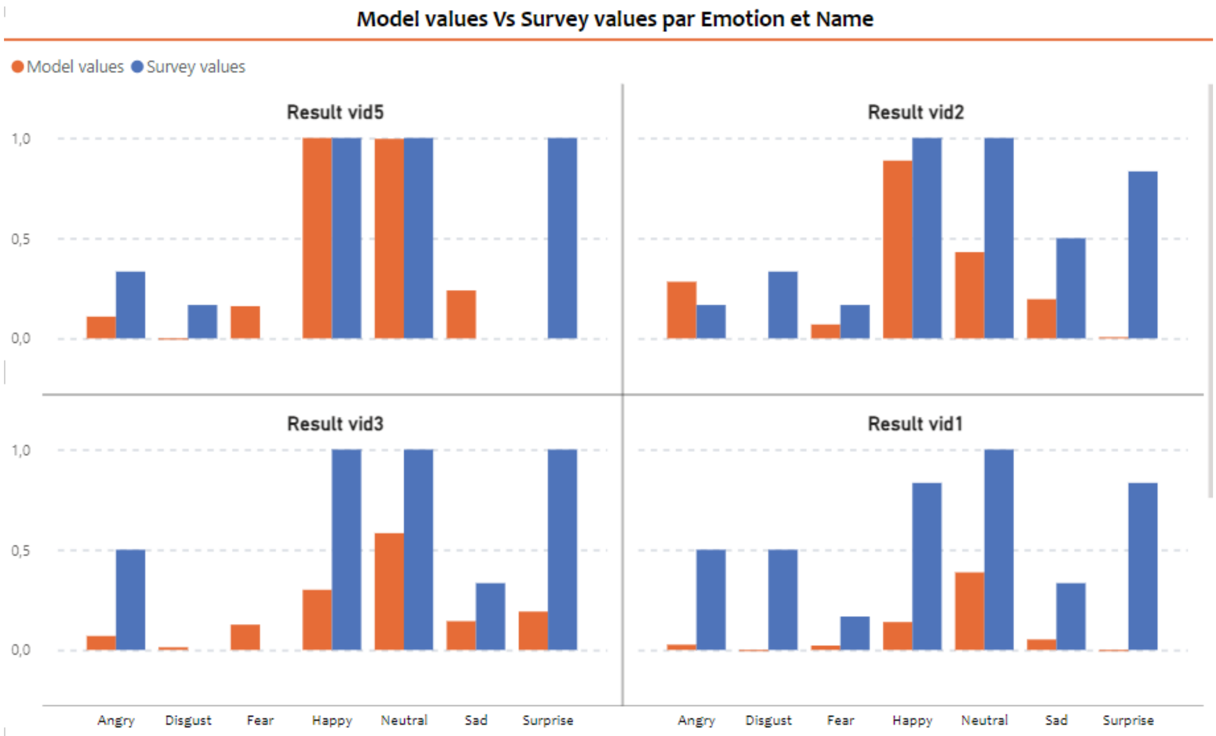


Figure 4.1 - Clustered column chart of survey vs model values 1.

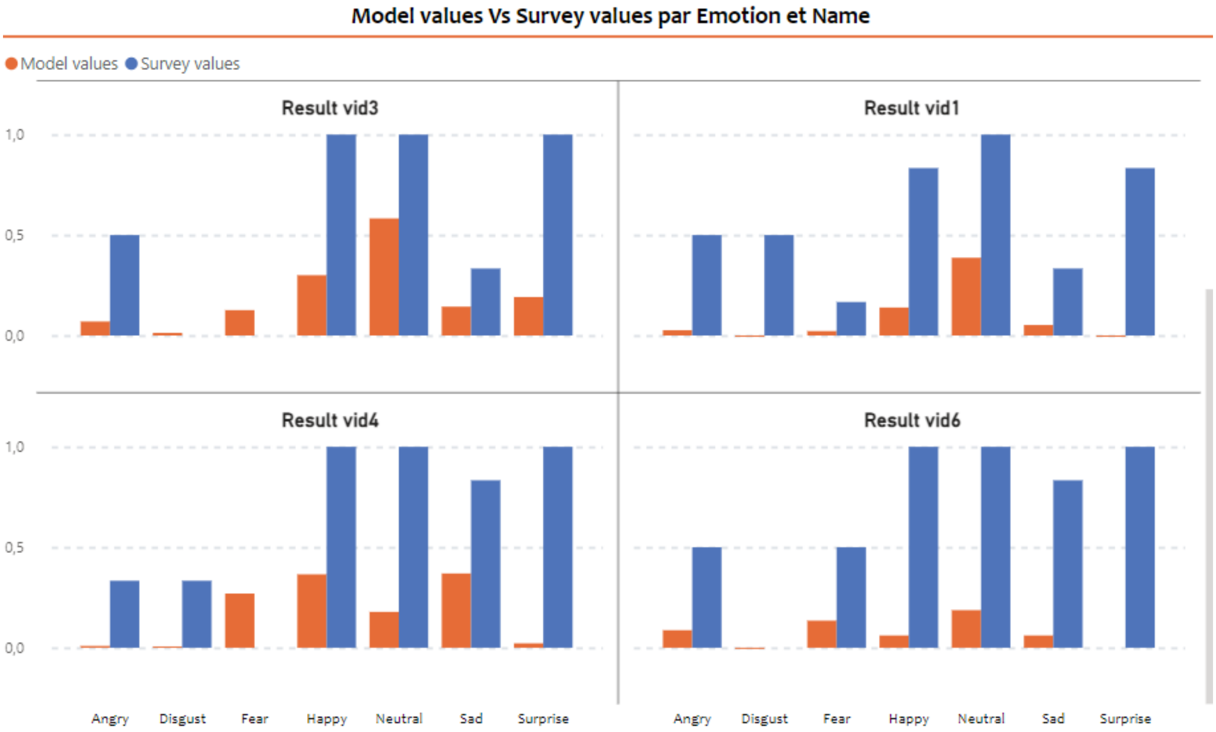


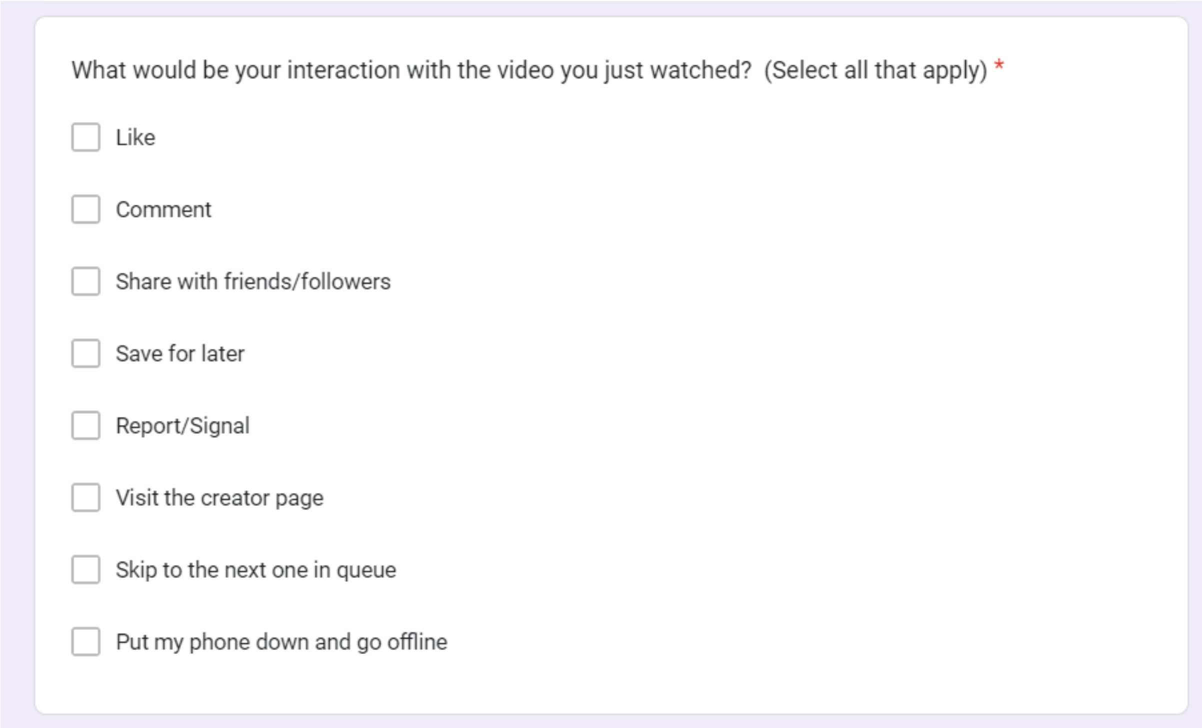
Figure 4.2 - Clustered column chart of survey vs model values 2.

In the first video, the top two emotions are the neutral and the happy from both data sources, which creates an alignment of their results. The difference is in the third emotion, where the model recorded Sad emotion and the survey recorded surprise. The second video also shows an alignment between the first two emotions, the happy and neutral ones. The difference is in the third emotion, where the model records angry as its top third and the survey records surprise as its top third. The third video shows an alignment between all the top three emotions, which are neutral, happy, and surprise. The fourth video shows an alignment in the top emotion recorded, the happy emotion. The difference is in the top second one, where the model predicted sad emotion and the survey recorded the neutral emotion. The top third for the model is fear, while in the survey it is the surprise. The fifth video insight is similar to the insights drawn from the first video. where the alignment exists between the happy and neutral emotions, while the third emotion differs between the model prediction of sad and the survey record of surprise. The sixth and last video record an alignment of the top emotion, neutral, and differs in the second and third emotions, in which the model predicts fear and angry respectively, while the survey shows happy and surprise, respectively.

Hence, through these insights, we can notice that the data collected from the survey always resulted in the top 3 emotions: happy, neutral and surprise. The difference lies in the fourth emotion in ranking, which varies from one video to another. While the model also predicts neutral and happy as the top emotions in most videos, a difference occurs in the third top emotion. This misalignment that occurs in the third emotion is also justified by the paper published by the researchers who built this model and found through its training an accuracy of 70% when it comes to emotion recognition.

Generally, both data sources indicate a joyous emotion dominating the videos posted by the content creator.

Next, we will take a look at the engagement indicators and their differences between the statistics pulled from the TikTok platform, also expressed in Table 3.1, and the engagement indicators that are collected through the survey responses for each video separately via the question in Figure 4.3.

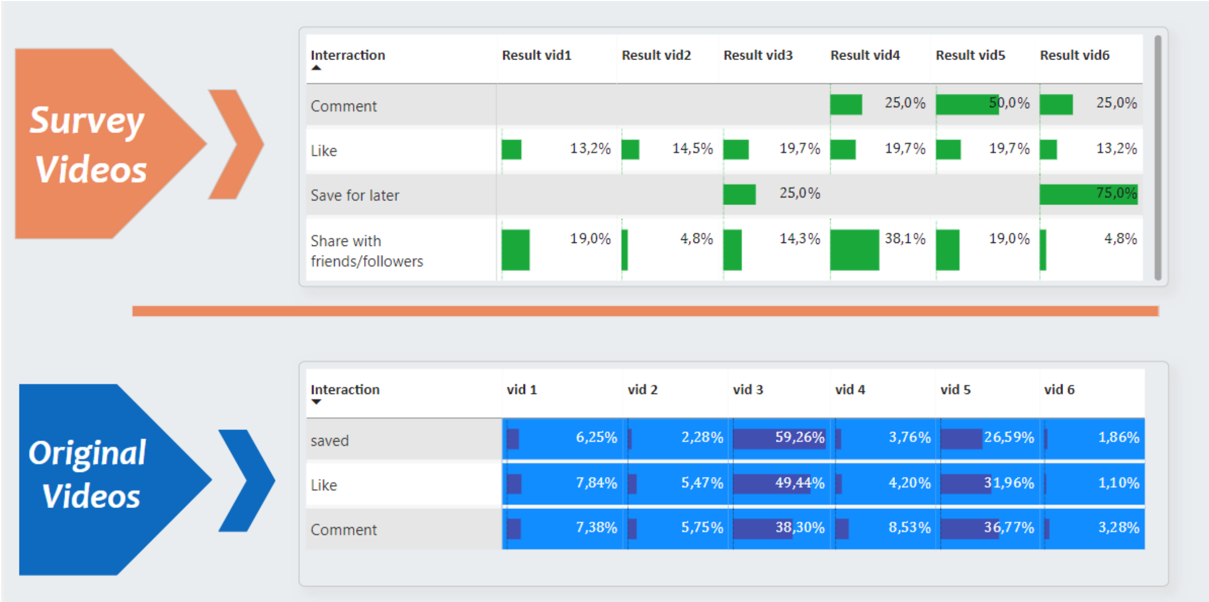

What would be your interaction with the video you just watched? (Select all that apply) *

- [ ] Like
- [ ] Comment
- [ ] Share with friends/followers
- [ ] Save for later
- [ ] Report/Signal
- [ ] Visit the creator page
- [ ] Skip to the next one in queue
- [ ] Put my phone down and go offline

Figure 4.3 - Engagement question in the survey.

From the analysis of this question, the top five engagements recorded were to like the video, followed by putting the phone down and going offline, sharing with friends and followers, skipping to the next one in queue, and visiting the creator page. However, from the platform statistics, only three measures of engagement could be gathered: like, comment, and save.

By putting both of these data tables in a matrix, as shown in figure 4.4, we can notice that there is a ranking that matches both videos in terms of liking the videos, which was the top engagement recorded from both tables. The highest from the survey was at 19.7%, while from the stats it was 49.44% for the third video, followed by the fifth video at 19.7% and 31.96%, respectively, and then there was a difference between the first and second videos in which the survey respondents liked them more than the platform users, with a difference of 1.3% from the survey table and 2.37% from the platform stats, which is a bit insignificant..



Interaction	Result vid1	Result vid2	Result vid3	Result vid4	Result vid5	Result vid6
Comment				25,0%	50,0%	25,0%
Like	13,2%	14,5%	19,7%	19,7%	19,7%	13,2%
Save for later			25,0%			75,0%
Share with friends/followers	19,0%	4,8%	14,3%	38,1%	19,0%	4,8%

Interaction	vid 1	vid 2	vid 3	vid 4	vid 5	vid 6
saved	6,25%	2,28%	59,26%	3,76%	26,59%	1,86%
Like	7,84%	5,47%	49,44%	4,20%	31,96%	1,10%
Comment	7,38%	5,75%	38,30%	8,53%	36,77%	3,28%

Figure 4.4 - Matrix of engagement metrics from the survey vs the platform.

What is worth mentioning in this matrix is the sharing with friends and followers' engagement metric recorded by the survey respondents. This metric, as we saw in the typology of ad sharers in the literature review, is one of the important behavioral responses that markers are looking for in every advertisement campaign. From the survey results, the video with the most shares was the fourth video, which was ranked 5th in likes from the platform stats. If we go back to the comparison of the emotions clustered in the column chart, we can notice that both happy and sad emotions are dominant in the survey records and the ML model prediction. This contrast in emotions shows an insight that it could be one of the driving factors of the quality engagement of the users. A video that contains both contrasting emotions could also mean that there could be an arc of emotions in the video. Either the video starts from a sad scenario into a happy ending, or the opposite starts from a happy scenario into a dramatic melancholic ending.

The second-most saved videos are both the first and the fifth. Looking back at the emotional comparison, we can see that if we eliminate the neutral emotion from the chart from both videos, we can see that the top three emotions from both videos would be happy, surprise, and angry. Two of these emotions are characterized by high intensity emotion as we have seen before from the Russel Circumplex model of the 2D emotional analysis. This proves that a video that conveys a message through intense emotions catches the attention of viewers and causes them to engage in the most desirable behavior, sharing. So, we can also say that an intense moment, facial expression, or video effect that could display an intense emotion is also a driving factor in favorable engagement.

Other insights into the driving factors of engagements could be gained from the demographic and behavioral analysis done through the survey data collected via a series of questions.

The most favorable time to consume social media content is during breaks and commuting, which is considered killing time. Both account for 79%. Thus, choosing a convenient time to display the marketing campaigns through social media should be carefully chosen and fall under one of these times: late morning, mid-day, or early evening. This will give more exposure to the content and contribute to its virality and engagement. The result is shown in Figure 4.5.

Time to consume	%
During Breaks (work/study breaks)	41,76%
During Commute Times	29,12%
Anytime I Get Notifications	14,29%
Free Moments	8,24%
Vacations	6,59%

Figure 4.5 - table represents times to consume media content.

Another metric that was recorded via the survey responses was the type of desirable content. Figure 4.6 shows the percentage of participants responses regarding the type of content. 82.42% of respondents are drawn to short, entertaining videos. This finding also supports the emotional responses collected and predicts that the happy and surprise emotions were present in all successful videos, as well as the degree of their intensity.

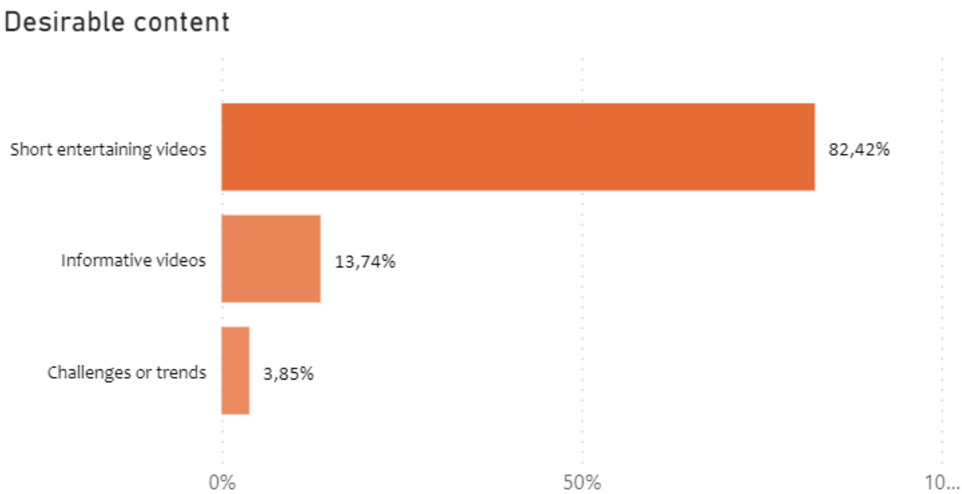


Figure 4.6 - stacked bar chart of desirable content.

The last metric measured was exposure to media content. Figure 4.7 demonstrates the ranking and percentages of the attributes and their exposure level. The dominant one is the recommendation of the platform. This attribute is hard to measure since it depends on the algorithm of the platform, which is not open source and changes across time. So, by omitting this choice, we are left with a dominant percentage of 12.63% that belongs to sharing methods both through instant messaging and the activity of their social network, compared to 6.59% that comes from following specific creators. This insight also proves that the sharing behavior discussed is the most favorable due to its ability to spread content, increase exposure, and result in a successful marketing campaign.

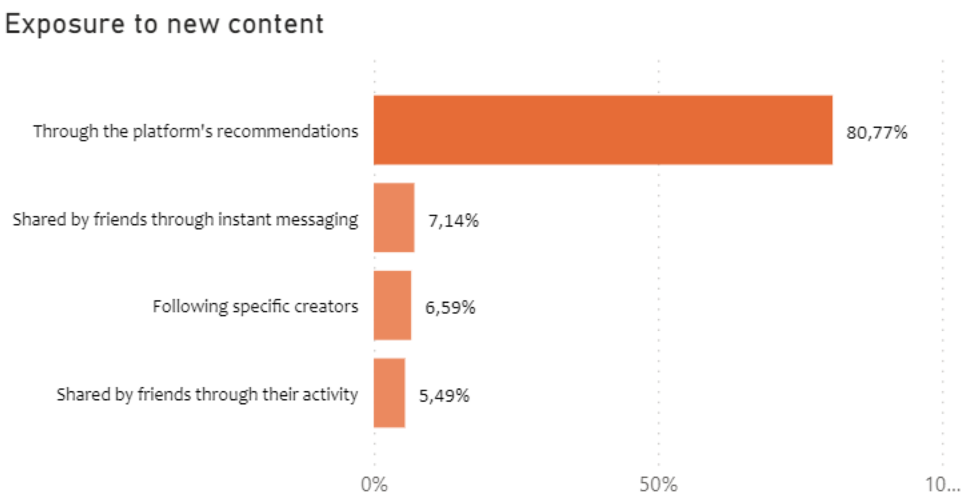


Figure 4.7 - stacked bar chart of exposure to new content.

We saw previously that the most desirable content recorded by users was entertaining videos, with a high percentage of 82%. Thus, in an attempt to automate the prediction of emotions so that marketers could run pilot tests of their campaigns and understand their emotional impact on engagement, Figures 4.8 and 4.9 outline the incentives for engagement that match the emotions detected by the model running through the test videos.

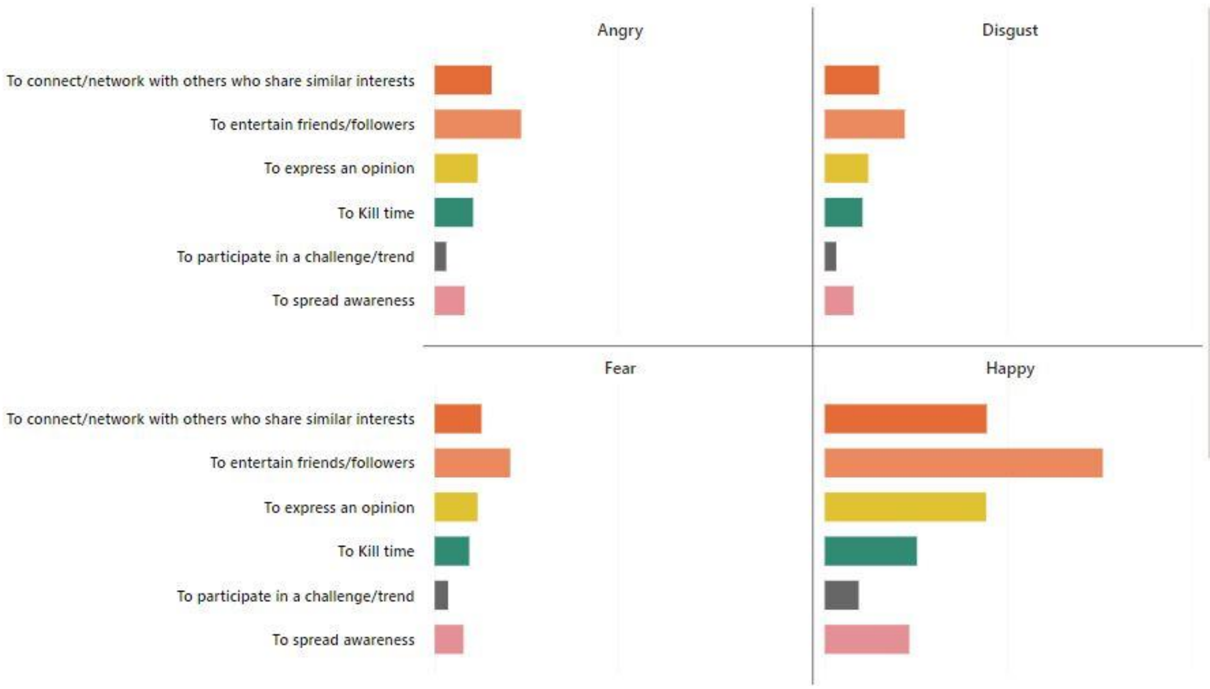


Figure 4.8 – Reasons of engagement by emotions predicted 1

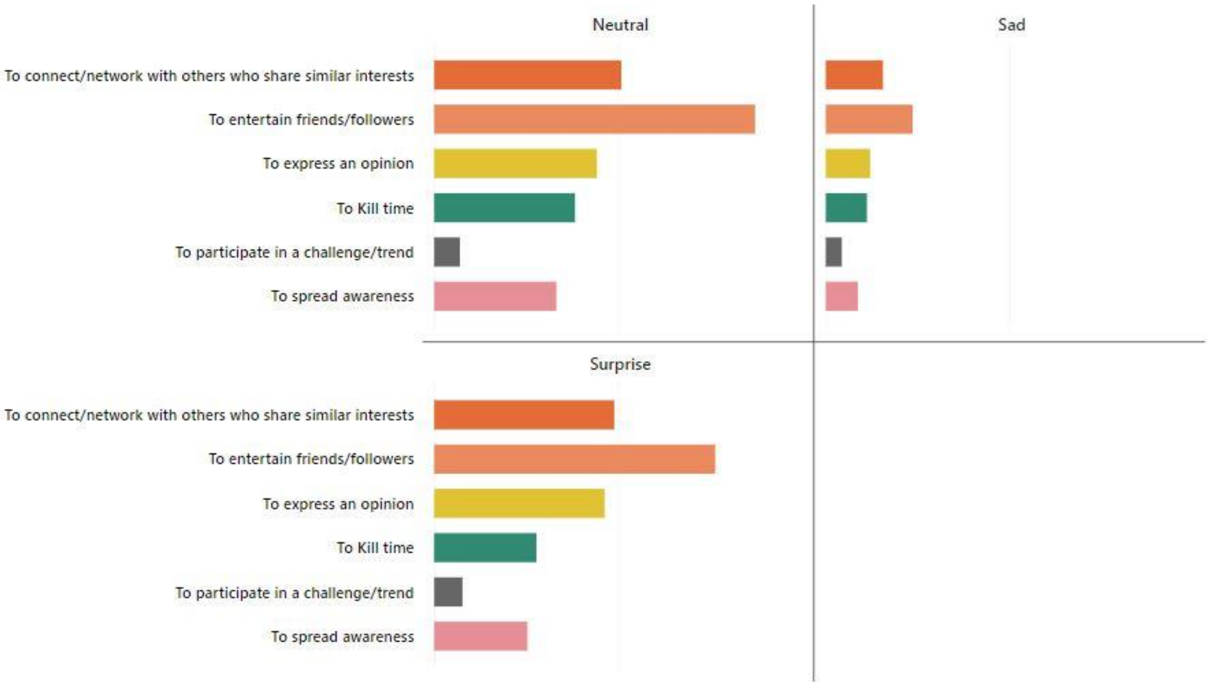


Figure 4.9 – Reasons of engagement by emotions predicted 2

The results show that actually entertaining friends or followers is the most sought-out reason to engage, as recorded through all the emotions predicted. However, we see a spike in this type of engagement in the Happy and the Surprise emotions. These two emotions were very dominant in the test videos, which can be justified by the niche of videos the content creator chosen for this test made. As a counterexample, the sixth video was supposed to raise awareness of a trailer for a new movie. Despite the fact that it had amazing audio and visual effects and the involvement of celebrities and influencers, it faced the lowest rates and numbers of engagement (refer to Table 3.1 for statistics). This could be interpreted as the content creator trying to raise awareness and drive this type of engagement by using the wrong emotions to support it in this video campaign, resulting in an engagement failure.

The second most desirable incentive for engagement is to connect or network with others who share similar interests. This was high in the happy and surprise emotions but also in the sad and angry emotions. Both of these emotions show different intensities of opposite valences. Hence, dramatizing emotions in short videos will push the users to engage by sharing with their network, resulting in a wildfire effect and desirable results for the marketing campaigns.

It is important to factor all of these nuances into the creation of a marketing campaign. The facial recognition and emotional prediction models developed and used by OpenCV, along with the data modeling and dashboarding, will create an ecosystem of software that applies sentiment analysis methodology and simulates user behavioral and emotional responses via an understanding of their emotional experience towards various video-centric marketing campaigns. Enabling an automated and dynamic process that facilitates the making and testing of marketing campaigns with an emphasis on the emotional aspect of them and the engagement they drive.

5. CONCLUSIONS AND FUTURE WORKS

Throughout this study, we found that the emotional response to video format media content can influence engagement rates and behaviors towards it. Thus, understanding the emotions generated by media content is essential in designing and testing a video-centric marketing campaign. This study used a sentiment analysis approach to perform a comparative analysis between Haar classifier emotion recognition models and survey-generated responses towards different video media contents. Then, from the predictions, modeling the data generated and visualizing the results using dashboarding software.

The emotions predicted from the facial detection and emotional recognition models show up in different ways in the test videos. One way of expressing presence is to have two opposing emotions that occur in the form of an arc, preferably telling a story, as in, start with one and finish with another in the opposite valence. This method of utilizing emotions triggers an engaging behavior in the viewer, and very often it is the sharing behavior that is most desirable by campaign leads. Another way is to dramatize the presence of the emotions by increasing their intensity in the video while playing with different parameters of content characteristics.

Other factors are also key in driving engagements, such as exposure to media content and choosing the right timing at which a large portion of users will be online and waiting to digest the content. An example of this would be during work or study breaks and in commute hours, as survey data shows. In addition to that, focus on the right type of content that users preferably consume and tend to share through their pages or instant messaging. The most dominant one from this study was entertainment, where the focus was on the positive and intense emotions, which were also recorded the most by our machine learning models and survey respondents.

5.1. THEORETICAL CONTRIBUTION AND BUSINESS IMPLICATIONS

The contribution of this study stems from its unique combination of sentiment analysis, survey insights, machine learning models, and a focus on optimizing marketing campaigns through personalized emotional engagement.

This study contributes to the existing literature by providing a new and different method to analyze social media videos using sentiment analysis. This study uniquely integrates sentiment and emotion analysis of digital content with survey data collected directly from users. While existing literature may focus on one of these aspects or use a triangulation method as an approach to their multimodal sentiment analysis towards media content (Bardzell, Bardzell, & Pace, 2009). The synthesis of both provides a more comprehensive understanding of user emotions and preferences.

By incorporating machine learning models for emotion prediction, this study extends the literature on emotion recognition in digital content. The comparison of model-predicted emotions with user-reported emotions provides valuable insights into the reliability and accuracy of such models, which could be explored further to improve their accuracy or find new ways to discover patterns and correlations between the content attributes of media content and the emotions it triggers.

Furthermore, this study has implications for designing, testing, and advertising marketing campaigns. Insights into how users discover and engage with content provide valuable information for improving content discovery methods. This can lead to enhancements in recommendation algorithms, search functionalities, and overall platform usability. Also, leveraging emotional analysis to inform marketing strategies provides a competitive edge. Brands that can connect with their audience on an emotional level are more likely to build lasting relationships and loyalty, differentiating themselves from competitors.

The study also promotes data-driven decision-making that helps marketers make informed decisions to optimize their campaigns. Understanding when users are most emotionally engaged with content allows for the scheduling of campaigns at optimal times. This can lead to higher visibility, as content is more likely to be noticed and shared during periods of peak emotional engagement. Marketing ethics are also implicated here. Avoiding content that evokes negative emotions or exploits emotions in an unethical manner is essential for maintaining brand trust and integrity.

5.2. LIMITATIONS AND FUTURE RESEARCH

Every research project comes with limitations that allow future researchers to exploit these gaps, build on top of the existing literature, and improve the quality of their work and insights.

One of the limitations of this study was the uncontrolled environment in which the survey was conducted. Due to the online nature of the survey, we could not determine or perform an initial test to assess the starting emotional condition of each participant, and thus an emotional baseline was hard to determine. Participants come from different emotional backgrounds, which could cloud their cognitive appraisal and influence the response they would give to the media content.

It is also hard to express emotions, especially when they are interrelated. In the case where I proposed the survey in person to a participant, a lack of understanding of their self-emotional state was very clear for certain media content. Which in turn causes randomness while choosing a score to give to the emotions.

Another limitation that was inevitable since the start of the study was the lack of resources, mainly the duration of the study and the choice of technologies. The use of commercial software comes with a lot of support from the manufacturer and is ready to deploy APIs and code, while open source lacks some documentation on certain specific issues. Then there is the heavy task of developing an algorithm that integrates the model into the local ecosystem and gets valid results out of it. The time constraint was also a noticeable limitation, since at the feasibility stage more data was projected to be collected from participants, which would allow more accurate results from the comparative analysis. In addition to finding dead ends with the coding part and the choice of technologies to be implemented.

This study is not comprehensive. The field of opinion mining is vast and keeps on growing at a fast pace with the rapid technological development that occurs. Regarding the continuity of this research, several points could be improved, such as:

Analyzing the prose reviews and aligning the results with the ML model, survey responses, and platform statistics. This question from the survey was intended for text analysis, so as to implement a multimodal approach to opinion mining (see Appendix B question 13 for reference). The idea was to analyze the images from the video frames and predict the emotions. Then, by using Python and the support vector machine algorithm (SVM), we analyze the text extracted from the prose reviews. The SVM method analyzes text by using a function that converts the words into vectors in numeric form.

The output is fitted into an SVM pipeline. (Rao, Ahuja, Kansara, & Patel, 2021) The emotional score would then be a ratio of text analyzed and video analyzed with a percentage that is based on their accuracy. Then average them based on the number of clips analyzed. This method will add a supporting modal to ensure the accuracy of the emotions predicted.

Another future work would be to test these results with different emotional recognition models and compare model results with survey results. So far, we only used the model developed by b-it Bonn-Aachen International Center for Information Technology and Hochschule Bonn-Rhein-Sieg. This model had an accuracy of 70%, which is rather high compared to other models but still low in terms of getting irrefutable results.

.

6. BIBLIOGRAPHICAL REFERENCES

Abraham Yousef. (2023, April 01). *The Latest Trends on Social Media Apps - Q1 2023*. Récupéré sur sensortower: https://sensortower.com/blog/the-latest-trends-on-social-media-apps-q1-2023

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L. P. (2018). OpenFace 2.0: Facial behavior analysis toolkit. In 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018) (pp. 59-66). IEEE

Bardzell, S., Bardzell, J., & Pace, T. (2009). Understanding Affective Interaction: Emotion, Engagement, and Internet Videos. *IEEE*, 1-8.

DOLAN, B. (2023, February 09). *what is Tiktok?* Récupéré sur investopedia: https://www.investopedia.com/what-is-tiktok-6826240

D'SOUZA, D. (2023, October 22). *TikTok: What It Is, How It Works, and Why It's Popular*. Récupéré sur investopedia: https://www.investopedia.com/what-is-tiktok-4588933#:~:text=TikTok%20allows%20users%20to%20make,the%2018%E2%80%9324%20age%20range.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 169-200.

Fresh Essays. (February 2023). Psychological Sensors for Facial Expression Recognition. Retrieved from https://samples.freshessays.com/psychological-sensors-for-facial-expression-recognition.html

Geyser, W. (2023, April 27). *The Ultimate Guide to Short-Form Video Content*. Récupéré sur influencermarketinghub: https://influencermarketinghub.com/short-form-video-content/

Gupta, S. (2022, 01 06). *How to apply Min-Max Normalization to your data*. Récupéré sur Medium: https://sourabharsh.medium.com/how-to-apply-min-max-normalization-to-your-data-f976d1633d2b

Gupta, Y., Agarwal, S., & Singh, P. (2020). TO STUDY THE IMPACT OF INSTAFAMOUS CELEBRITIES ON CONSUMER BUYING BEHAVIOR. Academy of Marketing Studies Journal Volume 24, 1528-2678.

Henkel, I. (2023, October 11). *Azure AI Video Indexer documentation.* Récupéré sur learn.microsoft: https://learn.microsoft.com/en-us/azure/azure-video-indexer/video-indexer-use-apis

Huang, M., Yang, Y., & Zhu, X. (2020). Short Video Understanding: Challenges, Methodology, and Applications. IEEE Transactions on Multimedia, 23, 6.

Kulkarni, K., Karlo, A., Sharma, D., & Sharma, P. (2019). A typology of viral ad sharers using sentiment analysis. *Journal of Retailing and Consumer Services*, 10.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining* . Morgan & Claypool Publishers.

Mittal, A. (2020, 12 20). *Haar Cascades, Explained*. Récupéré sur Medium: https://medium.com/analytics-vidhya/haar-cascades-explained-

38210e57970d#:~:text=A%20Haar%20classifier%2C%20or%20a,in%20an%20image%20and%20video

Octavio, A., Ploger, P. G., & Matias, V. (2019). *Real-time Convolutional Neural Networks for Emotion and Gender Classification.* Bruges: European Symposium on Artificial Neural Networks, Computational Intelligence.

Oliveira, T., Araujo, B., & Tam, C. (2020). Why do people share their travel experiences on social media? *Tourism Management 78*, 1-14.

OpenCV. (2023, 10 30). *Cascade Classifier*. Récupéré sur docs.opencv: https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html

OpenCV team. (2023). *About*. Récupéré sur opencv: https://opencv.org/about/

Palalic, R., Ramadani, V., Gilani, M., Gerguri-Rashiti, S., & Dana, L.–P. (2021). Social media and consumer buying behavior decision: what entrepreneurs should know? *Management Decision*, 1249-1270.

Poria, S., Cambria, E., Hazarika, D., Majumder, N., Zadeh, A., & Morency, L. P. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. Information Fusion, 37, 98-125.

Posner, J., Russel, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 715-734.

Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In Image Processing. 2002. Proceedings. 2002 International Conference on, volume 1, pages I–900. IEEE, 2002.

Rambocas, M., & Pacheco, B. (2018). Online sentiment analysis in marketing research: a review. *Journal of Research in Interactive Marketing*, 146-163.

Rao, A., Ahuja, A., Kansara, S., & Patel, V. (2021). Sentiment Analysis on User-generated Video, Audio and Text. *International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (p. 5). Mumbai: IEEE Xplore.

Rodriguez, A. (2022, December 01). *Internet illusionist Zach King breaks record for most viewed video on TikTok*. Récupéré sur guinnessworldrecords: https://www.guinnessworldrecords.com/news/2022/12/internet-illusionist-zach-king-breaks-record-for-most-viewed-video-on-tiktok-724834

SAMBARE, M. (2020, 10 16). *FER-2013*. Récupéré sur kaggle: https://www.kaggle.com/datasets/msambare/fer2013/data

Schreiner , M., Fischer, T., & Riedl, R. (2019). Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda. *Electronic Commerce Research*, 17.

Singh, A. (2023, August 7). *The Rise of Short-Form Video Content: Reshaping Digital Media Consumption*. Récupéré sur Linkedin: https://www.linkedin.com/pulse/rise-short-form-video-content-reshaping-digital-media-aastha-singh/

Viola , P., & Jones, M. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 137–154.

West, T. (2011). Going Viral: Factors That Lead Videos to Become Internet Phenomena. *Broadcast Journalism and International Studies*, 76-84.

Zadeh, A., Poria, S., Cambria, E., & Hussain, A. (2018). Multimodal sentiment analysis in the wild. In Multimodal Behavior Analysis in the Wild (pp. 135-155). Springer.

7. APPENDIX

7.1. APPENDIX A

Consent form to participate in the study

We are interested in understanding your emotional response to social media posts. You will be asked to answer some contextual questions, watch some short TikTok videos, then answer some questions about your experience with the videos watched.

Rest assured that your response will be treated with complete confidentiality and anonymity, and it will be used only for academic purposes in a synthesized format. At the end of the study, all collected data will be destroyed.

This survey should take you around 7 to 10 minutes to complete, depending on how well you can express your emotional experience.

All the content in this survey is written in English.

By clicking next, you acknowledge that your participation in the study is voluntary, and that you have the right to withdraw from this survey at any moment of time.

For any inquiries, please contact:

m20210139@novaims.unl.pt

7.2. APPENDIX B:

Survey form for Study

1. Please specify your age

2. Please specify your gender
   Male
   Female
   Others

3. Which country are you from?

4. How many hours a day you spend on social media?
   less than an hour
   between 1 to 3 hours
   more than 3 hours

5. What is your favourite medium to consume social media content
   Phone
   Tablet
   Computer
   Other

6. What is your preferred time to consume social media content? (Select all that apply)
   During Commute Times
   During Breaks (work/study breaks)
   Anytime I Get Notifications
   On Weekends
   Vacations
   During Creative Breaks
   When Seeking Information
   Free Moments
   Other

7. Which ways do you use to discover new content? (Select all that apply)
   Through the platform's recommendations
   Shared by friends through their activity
   Shared by friends through instant messaging
   Following Hashtags or trends
   Following specific creators
   Actively searching for specific content
   Other

8. What motivates you to engage with a piece of content on social media? (Select all that apply)
   To entertain friends/followers
   To express an opinion
   To spread awareness
   To participate in a challenge/trend

To connect/network with others who share similar interests
To Kill time
Other

9. What types of content do you prefer on social media? (Select all that apply)
   Short entertaining videos
   Informative videos
   Inspirational videos
   Personal vlogs
   Challenges or trends
   Promotions, sales, discounts, and giveaways
   Other

10. Which content attributes are likely to capture your attention when browsing your feed? (Select all that apply)
    Captivating thumbnail
    Attention-grabbing caption
    Short video duration
    Clear audio quality
    High visual quality
    Celebrity involvement
    The way it is positioned in your feed
    Number of views it got
    Recency of the video
    Other

11. What would be your interaction with the video you just watched? (Select all that apply)
    Like
    Comment
    Share with friends/followers
    Save for later
    Report/Signal
    Visit the creator page
    Skip to the next one in queue
    Put my phone down and go offline

12. After viewing the above video, please rate the extent to which you experienced the following emotions on a scale of 1 to 7 during your viewing session. 1 is strongly disagree and 7 is strongly agree.

	1 (Strongly disagree)	2	3	4	5	6	7 (Strongly agree)
Angry							
Disgust							
Fear							
Happy							
Neutral							
Sad							
Surprise							
Angry							
Disgust							
Fear							
Happy							

Neutral
Sad
Surprise

13. Please write a short informal review regarding each video (it could be one word or more)

14. Which video appealed to you the most out of the videos suggested?
    No title
    When the day needs me anywhere and everywhere, #CeraVe SPF is what I wear so I can
    This is Puzzling
    Can someone please explain how the mirror knows there is an egg?
    Life isnt always as it seems @sofyank96
    Spent the last 6 months making this, hope you enjoy the trailer. Watch the entire short film in my profile

15. What was the characteristic(s) that made you choose that video? (Select all that apply)
    Topic of the video
    Components of the video (such as pictures, links, emojis, hashtags)
    Duration of the video
    Interactivity of the video
    Originality of the content
    The order of the video watched
    Other

7.3. Appendix C:

Ethics commitee approval:

This is to certify that

Project No.: INFSYS2023-11-263148

Project Title: Factors driving engagement in social media – An analysis of the cognitive appraisal on video content

Principal Researcher: Younes Abouljid

according to the regulations of the Ethics Committee of NOVA IMS and MagIC Research Center this project was considered to meet the requirements of the NOVA IMS Internal Review Board, being considered APPROVED on 11/26/2023.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.

The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of

- Any significant change to the project and the reason for that change;
- Any unforeseen events or unexpected developments that merit notification;
- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 11/26/2023

NOVA IMS Ethics Committee

ethicscommittee@novaims.unl.p

8. ANNEXES (OPTIONAL)

Link to Tik Tok video used in the dataset by the survey and model.

Video 1:

https://www.tiktok.com/@zachking/video/7218568693398162734?is_from_webapp=1&web_id=7201136050294408710

Video 2:

https://www.tiktok.com/@zachking/video/7219102627735342382?is_from_webapp=1&sender_device=pc&web_id=7201136050294408710

Video 3:
https://www.tiktok.com/@zachking/video/7221922451012324650?is_from_webapp=1&sender_device=pc&web_id=7201136050294408710

Video 4:

https://www.tiktok.com/@zachking/video/7223843384174136618?is_from_webapp=1&web_id=7201136050294408710

Video 5:

https://www.tiktok.com/@zachking/video/7226754712283041070?is_from_webapp=1&web_id=7201136050294408710

Video 6:

https://www.tiktok.com/@zachking/video/7227146522977111338?is_from_webapp=1&web_id=7201136050294408710