



MVIndEmo: a dataset for micro video public-induced emotion prediction on social media

Zhenhua Guo¹ · Qi Jia¹ · Baoyu Fan¹ · Di Wang¹ · Cong Xu¹ · Yanwei Wang¹ · Yaqian Zhao¹ · Rengang Li^{1,2}

Received: 26 March 2023 / Accepted: 8 December 2023 / Published online: 4 February 2024
© The Author(s) 2024

Abstract

Distinct from the realm of perceived emotion research, induced emotion pertains to the emotional responses engendered within content consumers. This facet has garnered considerable attention and finds extensive application in the analysis of public social media. However, the advent of micro videos presents unique challenges when attempting to discern the induced emotional patterns exhibited by content consumers, owing to their free-style representation and other factors. Consequently, we have put forth two novel tasks concerning the recognition of public-induced emotion on micro videos: emotion polarity and emotion classification. Additionally, we have introduced a accessible dataset specifically tailored for the analysis of public-induced emotion on micro videos. The data corpus has been meticulously collected from Tiktok, a burgeoning social media platform renowned for its trendsetting content. To construct the dataset, we have selected eight captivating topics that elicit vibrant social discussions. In devising our label generation strategy, we have employed an automated approach characterized by the fusion of multiple expert models. This strategy incorporates a confidence measure method that relies on three distinct models for effectively aggregating user comments. To accommodate adaptable benchmark configurations, we provide both binary classification labels and probability distribution labels. The dataset encompasses a vast collection of 7,153 labeled micro videos. We have undertaken an extensive statistical analysis of the dataset to provide a comprehensive overview composition. It is our earnest aspiration that this dataset will serve as a catalyst for pioneering research avenues in the analysis of emotional patterns and the understanding of multi-modal information.

Keywords Induced emotion · Emotion classification · Emotion polarity · Micro video · Social media · Multi-modal

1 Introduction

Induced emotion, distinct from perceived emotion that convey from content creators, refers to the emotional reactions triggered in content consumers. Currently, there is a growing interest in understanding the patterns of induced emotion

Communicated by B. Bao.

✉ Qi Jia
jiaqi01@ieisystem.com

Zhenhua Guo
guozhenhua@ieisystem.com

Baoyu Fan
fanbaoyu@ieisystem.com

Di Wang
wangdi11@ieisystem.com

Cong Xu
xucong@ieisystem.com

Yanwei Wang
wangyanwei@ieisystem.com

Yaqian Zhao
zhaoyaqian@ieisystem.com

Rengang Li
hsslab@gmail.com; lirengang@gmail.com

- ¹ Inspur Electronic Information Industry Co.,Ltd., Beijing 100084, China
- ² Tsinghua University, Shuang Qing Road, Hai Dian, Beijing 100084, China

in various forms of emotional content, such as music [1, 2] and movie [3]. Meanwhile, induced emotion has found widespread applications in the analysis of public social media, spanning from measuring the effectiveness of commercial advertisements [4] to predicting public policies for social stability [5], among others. However, the emergence of micro video, which have gained immense popularity and wield unprecedented influence on social media platforms, has prompted individuals to express and exert their influence more vividly compared to traditional text and image-based content. Consequently, studying the public-induced emotion patterns, particularly in the context of micro videos, has become indispensable.

In contrast to previous research, analyzing the public-induced emotion on micro videos presents distinct challenges. Firstly, while previous studies have often focused on professional creations such as music and movies, micro videos showcase a free-style representation that can be created by anyone. This freedom in content creation poses difficulties in understanding and analyzing the induced emotion patterns on micro videos. Secondly, micro videos lack the synchronized subtitle text that is typically present in movies, which has been shown in previous research to play a crucial role in recognizing induced emotions. Instead, micro videos often have short text descriptions that summarize their content but do not provide synchronized textual information alongside the video. Lastly, our proposed task specifically targets the public aspect of induced emotion. Unlike analyzing individual emotions, the induced emotions observed in the public can be complex, diverse, and even incompatible due to cultural and educational variations among people. Therefore, possessing common knowledge about social dynamics becomes crucial in this context. Furthermore, traditional one-hot classification patterns may not adequately capture the nuances of public-induced emotion recognition. Instead, employing a probability distribution representation

proves to be more effective, which motivates researchers to develop flexible benchmark settings for this task.

To promote research on public-induced emotion patterns in micro videos and address the challenges, we have developed two public-induced emotion recognition tasks: emotion polarity [10] and emotion classification [11]. These tasks find wide-ranging applications in the realm of social media [4, 5]. To facilitate these tasks, we have designed an automated method for constructing a micro video public-induced emotion recognition dataset. In our approach, we infer the public-induced emotion patterns in micro videos by aggregating the emotions and polarities expressed in user comments. We have incorporated a multi-expert fusion strategy based on the confidence measure method to minimize inference errors. For each task, we provide two forms of labels: binary classification labels and soft labels that reflect the probability distribution of public-induced emotions. This setting aligns with the realistic nature of public-induced emotion patterns and proves useful in establishing more flexible benchmarks, such as emotion regression, ranking, and retrieval. We illustrate the differences between our dataset and existing datasets focused on induced emotion in Table 1. To the best of our knowledge, this dataset represents the first public-induced emotion recognition dataset specifically tailored for micro videos. The key contributions of our work can be summarized as follows:

The main contributions of our work can be summarized as follows:

1. We introduce two tasks for public-induced emotion recognition in micro videos: emotion polarity and emotion classification. Additionally, we construct the micro video public-induced emotion recognition dataset, which focuses on the multi-modal nature of induced emotion patterns in social media.
2. We carefully select topics that are currently capturing social attention and develop an automated method to

Table 1 The summary of the previous dataset about induced emotions to reveal the difference

Dataset	Domain	Modal	External feature	Personal/Public	Polarity	Class	V/A
emotifydata [1]	Music	a	–	Personal	No	Binary	No
emotion-recognition [2]	Music	a	–	Personal	No	Binary	Yes
LIRIS-ACCEDE [3]	Movie	v, a, t	Yes	Personal	No	No	Yes
Emotion6 [6]	Flickr	v	–	Personal	No	Binary, Distribution	Yes
IESN [7]	Flickr	v	Yes	Personal	No	Binary	Yes
RETWEET [8]	Twitter	t	–	Public	Binary	No	No
ReactionGIF [9]	Twitter	t	–	Personal	No	Binary	No
Our Dataset	Micro Video	v, a, t	–	Public	Binary, Distribution	Binary, Distribution	No

For the column of Modal, v, a, t means visual, audio, text, respectively. The column of External Feature indicate that the related research whether use other features except the raw data. The column of Personal/Public represent the object of dataset. There are three kinds task setting: emotion polarity, emotion classification and Valence-Arousal score (V/A). For polarity and classification, there are two label schema: Binary and possibility distribution

generate public-induced emotion labels using user comments. To improve the accuracy of inference, we employ a multi-expert fusion strategy based on the confidence measure method.

3. For both recognition tasks, we provide binary classification labels as well as probability distribution labels. This enables more flexible benchmark settings, such as emotion regression, ranking, and retrieval tasks.

We believe that this dataset will inspire new research directions in sentiment analysis, particularly in the area of induced emotion, and provide valuable support to scholars studying social media and multi-modal information understanding.

2 Related work

In this section, we aim to explore the existing research on induced emotion patterns, including datasets and methods. We begin by distinguishing between perceived emotion and induced emotion from the perspective of information reception. Perceived emotion refers to the emotion expression identified from the content of the received information, while induced emotion refers to the emotion evoked by the stimulation of the received information [12, 13]. In many previous studies on sentiment analysis, the distinction between perceived emotion and induced emotion has not been clearly defined. However, some research has demonstrated that perceived emotion is not always consistent with induced emotion [13]. Due to the powerful dissemination capabilities of micro video social media platforms, the analysis of induced emotion becomes more crucial than perceived emotion. Moreover, it has a wide range of applications in various perspectives [4–6, 14]. Therefore, it is essential to focus on induced emotion analysis, considering the significant impact and diverse applications in the context of micro video social media.

Indeed, while the research on induced emotion has received less attention compared to perceived emotion, it has seen significant development in recent years. In a notable early work, Kallinen et al. [12] introduced the distinction between perceived emotion and evoked emotion, focusing on the realm of music. Aljanaki et al. [1] explicitly proposed a dataset for studying induced emotion recognition in music. They utilized the GEMS music emotion model to define the prediction target and collected induced emotion labels through a game-like approach. Song et al. [2] explored the difference between perceived emotion and induced emotion in Western popular music. They approached the topic from the perspective of emotion classification and dimensional models. The study released a relevant dataset and discussed the influence of individual differences on

emotion judgments. It is worth noting that research on induced emotion in the domain of music requires high-quality, clean, and noise-free audio data due to its impact on emotion perception and elicitation. This aspect underscores the importance of ensuring the audio data used in such studies meets specific standards and requirements.

Induced emotion patterns in the movie field have also started to gain attention, with researchers exploring the relationship between movie content and audience emotional response. Benini et al. [15] argued that movie scenes sharing similar connotations could evoke the same emotional response. They proposed an approach to infer induced emotions of the audience by establishing a mapping between the connotative space and emotional categories. Tian et al. [13] made a significant contribution by distinguishing between perceived emotion and induced emotion in the context of movies. They investigated the relationship between these two types of emotions and used multiple modalities for feature extraction. This included audience physiological signals recorded using wearable devices, as well as audio and visual features, lexical features from movie scripts, dialogue features, and aesthetic features. Their study used an LSTM-based model to recognize induced emotions, demonstrating the effectiveness of fusing multiple modalities. The research findings highlighted the inconsistency between perceived emotion and induced emotion in the audience. A subsequent work by Muszyński et al. [16] further explored the relationship between dialogue and aesthetic features in induced emotions in movies. They proposed a novel multi-modal model for predicting induced emotions. Liu et al. [17] took a different approach by using EEG signals to infer induced emotions of the audience in real-time while watching movies. Their study focused on the LIRIS-ACCEDE database [3], which has been widely used in recent research on movie induced emotions. The primary goal of these studies was to advance research in movie art and assist filmmakers in designing emotionally engaging content. However, some challenges remain. The extensive and costly nature of modalities like EEG signals poses limitations on their widespread adoption. Moreover, the research in this domain has mostly concentrated on movies with dialogue-driven plots, which narrows the scope of the research topics.

Recently, there has been a growing interest in recognizing induced emotions in the context of social media. The initial research in this area can be traced back to the SemEval-2007 Tasks [18, 19], which focused on how news headlines evoke emotions in readers. The early studies primarily revolved around emotion lexicons and topic mining. Bao et al. [20, 21] proposed emotion topic models to predict readers' emotional responses to news topics. They utilized the Facebook Reactions function to construct a dataset and set emotional targets for prediction. Rao et al. [22] also explored affective topic mining in the context of news articles. Peng et al. [6]

focuses on induced emotions in viewers based on photographs. They introduced a CNN-based model for recognizing induced emotions. Recognizing the complexity of human emotions, they employed induced emotion distribution as the target variable rather than simple classification. They further discussed the effect of altering features like tone and texture of a photo on induced emotions. Clos et al. [23] designed two lexicon-based baseline methods for recognizing readers' induced emotions toward news articles posted on the Facebook pages of The New York Times. Zhao et al. [7] published a dataset aiming to examine the induced emotions of individuals toward pictures posted on Flickr. The paper emphasized the presence of individual differences, highlighting that different people may have different induced emotions toward the same image. The dataset not only considered the content of the image but also included the social context of the user. These studies demonstrate the increasing recognition of induced emotions in the context of social media.

3 Materials and methods

Our dataset originates from TikTok, serving as the primary source for our data collection pipeline. In this section, we will elaborate on the process of our dataset collection

pipeline, shed light on our labeling strategy, and present crucial analysis statistics derived from the dataset.

3.1 Overview of Tiktok Website

TikTok is an emerging social media platform that has gained significant popularity for its micro video content creation and sharing. While primarily a mobile application, TikTok also offers a website interface, as depicted in Fig. 1. Each micro video page on TikTok comprises a user-generated video, accompanied by a descriptive caption enriched with relevant hashtags. Additionally, users can express their thoughts and engage in discussions through comments, and demonstrate appreciation by liking specific comments. Similar to Twitter, TikTok employs hashtags to categorize micro videos based on various topics like **#football**, **#economy**, and more. Consequently, we leverage these hashtags to collect a dataset consisting of micro videos pertaining to various social themes.

3.2 Data collection pipeline

We have devised a distinctive data collection pipeline to accommodate the framework of TikTok. This pipeline encompasses two crucial stages: (1) the collection of micro videos and accompanying comments through the

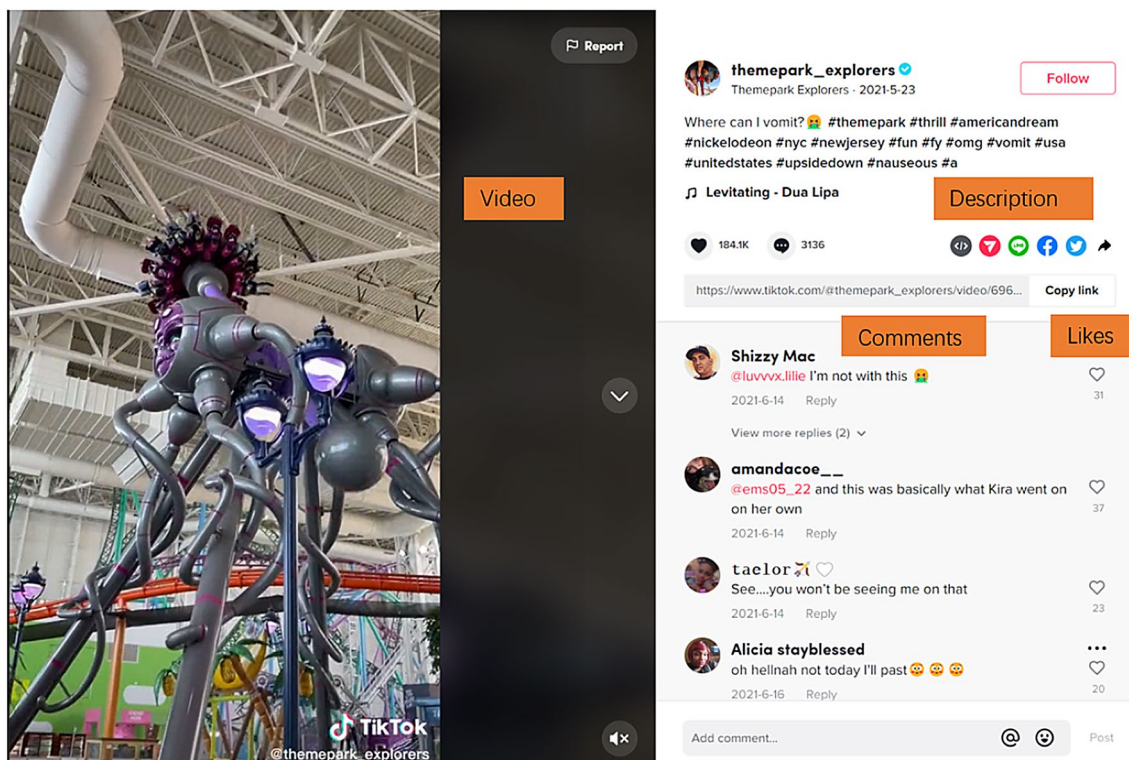


Fig. 1 Micro video page in Tiktok website. The key information of micro video are identified with orange text boxes in the figure

utilization of relevant hashtags; (2) the meticulous process of cleansing these micro videos and comments.

In order to gather micro videos and comments, we employ a methodical approach of collecting raw data by utilizing carefully selected hashtags. We conduct extensive research by perusing reputable news outlets such as CNN, New York Times, and FOX News to identify noteworthy topics that have recently captured public attention. Subsequently, we employ targeted keyword searches on the TikTok website to identify the relevant hashtags associated with these topics. To guide our data collection efforts, we have identified eight primary topics, each corresponding to a range of hashtags as depicted in Table 2. These hashtags, generated organically by users, serve as indicators of the prevailing subjects currently being discussed on social media. By accessing the hashtag pages, we obtain pertinent micro video content along with its accompanying textual descriptions. Simultaneously, we capture the corresponding comments that accompany each video. It is important to note that we exclusively collect first-generation reply comments for each video, disregarding further sub-level responses to maintain focus and streamline our data collection process.

To ensure the representativeness of the micro videos and the meaningfulness of the comments, a rigorous data cleaning process is implemented. Initially, micro videos with fewer than 1,000 comments are eliminated from the dataset. This step guarantees that the selected micro videos have achieved a substantial level of visibility and have garnered public attention on TikTok. Given that TikTok is a global social platform, it is noteworthy that micro videos may receive comments in various languages. In order to mitigate potential cultural differences arising from diverse language backgrounds, we exclusively utilize English comments as the foundation for generating induced emotion labels. This allows for a consistent and standardized approach in capturing the underlying emotional sentiments. Furthermore, certain comments serve specific purposes such as attracting attention from other users, often in the form of mentioning a particular user or incorporating hashtags to expand upon the topic being discussed. In order to streamline the

dataset, comments containing patterns like "USER" or "#HASHTAG" are eliminated. However, it is important to highlight that the number of likes assigned to each comment is preserved as it holds substantial importance and serves as a valuable metric for assigning weightage during the label generation process.

3.3 Label strategy

In order to capture the collective public induce emotion responses, it is imperative to incorporate the sentiments expressed in social media comments. Our approach revolves around analyzing the sentiment of these comments and subsequently generating a comprehensive understanding of the public-induced emotions on the micro videos. To accomplish this, we propose leveraging existing sentiment analysis tools to analyze the sentiment conveyed in the comments. In order to mitigate any potential biases stemming from relying solely on a single tool, we employ a combination of three distinct methods for analyzing the emotional content of each comment. Subsequently, we integrate the inference results from these methods to derive a holistic understanding of the public-induced emotions. We utilize the sentiment analysis method provided by the Huggingface model library, as outlined in Table 3, to facilitate this process. By incorporating multiple sentiment analysis techniques, we strive to enhance the reliability and accuracy of our predictions.

Table 3 The sentiment analysis method for each tasks

Task	Method name
Polarity	monologg/bert-base-cased-goemotions-group cardiffnlp/twitter-xml-roberta-base-sentiment finiteautomata/bertweet-base-sentiment-analysis
Classification	trnt/twitter_emotions bhadresh-savani/roberta-base-emotion bhadresh-savani/electra-base-emotion

These methods are easier to obtain in Huggingface model library

Table 2 The social topics for data collection on Tiktok

Topic	Hashtag
Amarica	America, USA, NYC, Americandream, Immigrant, Mexican
Politics	Geopolitics, Ukraine, Ukrainear, WW3
Digital Currency	Bitcoin, BTC, Cryptocurrency, ETH, NFT
Sports	Football, Soccer, NBA, NFL
Olympics	Olympics, Athlete, Tokyo2020, Winterolympics, Teamusa
Corona Virus	Coronavirus, Happyathome, Lockdown
Economy	Economy, Gasprices, Inflation, Money, Stocks
Abortion	Abortion, Feminism, Prochoice, Prolife, Roewwade

Each topic involves several hashtags

For the emotion polarity task, each sentiment analysis method generates a probability vector that indicates the distribution of probabilities for the categories of **positive**, **negative**, and **neutral** emotions. These probability vectors provide insights into the emotional polarity conveyed by the comments. Similarly, for the emotion classification task, we follow the same approach as outlined in the Emotion dataset provided by Huggingface (<https://huggingface.co/datasets/emotion>). Each sentiment analysis method maps the comments into one of six predefined emotion categories: **anger**, **fear**, **joy**, **love**, **sadness**, and **surprise**. The output for this task is also in the form of a probability distribution vector that represents the predicted probabilities for each emotion category, capturing the emotional classification of the comments. These probability vectors provide insights into the emotions conveyed by the comments. For a comment of micro video, three models for each task generate the probability distribution vectors, namely R_1 , R_2 , and R_3 . These need to be fused to derive a comprehensive inference probability distribution. From an inference standpoint, the highest probability within the result signifies the most probable classification outcome. As such, we encounter three possible scenarios with the three methods:

Case 1: The classification outcomes across all three methods align.

Case 2: The classification outcomes of all three methods do not align.

Case 3: Among the three methods, two methods yield the same classification outcomes.

We have conducted an analysis on the consistency of inference outcomes for each task based on the three methods,

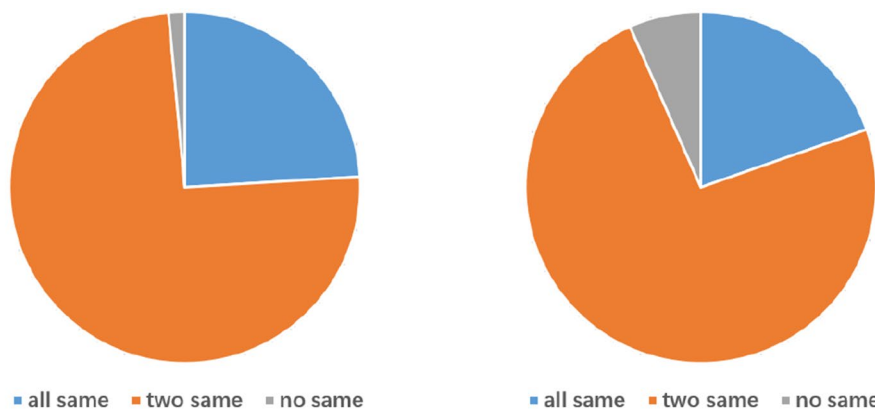
as displayed in Fig. 2. Notably, the inference consistency for polarity is superior compared to the classification task. In general, the proportion of **Case 2** is comparatively low. This serves as a reliable indicator for the accuracy of comment inference results.

In **Case 1**, we aggregate the three probability distribution vectors and subsequently normalize them to obtain the final inference result for a comment. However, for the other cases, it becomes necessary to devise a more rational fusion strategy.

Confidence plays a crucial role in determining the certainty of model inference results [24]. Shannon entropy is a commonly used measure for assessing the confidence of model predictions [25, 26], and it finds widespread application in model optimization and sample denoising [27]. The Shannon entropy is calculated for the probability distribution vector of the inference result. A lower entropy value indicates a more credible inference result, whereas a higher entropy signifies a less reliable inference result. The range of Shannon entropy values is determined by the length of the probability distribution vector, i.e., $0 \leq H(R_i) \leq \log N$, where N represents the length. Exploiting this concept, we employ Shannon entropy to fuse the three methods in order to derive the final comment inference results. In **Case 2**, we compute the Shannon entropy H or each method's inference results separately. To assess the confidence score of the inference result, we employ the following formula:

$$ConfidenceScore(H) = 1 - H / \log N. \tag{1}$$

Fig. 2 The consistency of the inference results of the three methods for each task. Figure **a** is about the polarity, and Figure **b** is about the classification



a) The distribution of prediction of polarity labels from three models

b) The distribution of prediction of emotion labels from three models

Algorithm 1 Label generation strategy

Input: input: all comments set C
Output: output: fusion label distribution L

```

1 for each comment in  $C$  do
2   calculate the probability distribution of three models  $R_1, R_2, R_3$ ;
3   get the label  $l_1^*, l_2^*, l_3^*$  corresponding to the maximum probability;
4   calculate the number of same label same_num;
5   if same_number=3 then
6     |  $R^f = \text{mean}(R_1 + R_2 + R_3)$  ;
7   end
8   if same_number=2 then
9     for each model  $j$  do
10      calculate information entropy  $H(X) = -\sum_{i=1}^n p(x_i) \log p(x_i)$ ;
11      calculate the Confidence  $c = 1 - H/\log N$ , where  $N$  is the
        dimensions of probability distribution;
12     end
13     denote the Confidence of same label models  $R_i^s$  as  $c_1^s$  and  $c_2^s$ ;
14     denote the Confidence of different label models  $R^d$  as  $c^d$ ;
15     if  $c_1^d > \text{Max}(c_1^s, c_2^s)$  then
16       |  $R^f = c_1^s \times R_1^s + c_2^s \times R_2^s + c_1^d \times R^d$ ;
17     else
18       |  $R^f = \text{mean}(c_1^s \times R_1^s + c_2^s \times R_2^s)$ 
19     end
20   end
21   if same_number=0 then
22     for each model  $j$  do
23       calculate information entropy  $H(X) = -\sum_{i=1}^n p(x_i) \log p(x_i)$ ;
24       calculate the Confidence  $c = 1 - H/\log N$ , where  $N$  is the
        dimensions of probability distribution;
25     end
26      $R^f = c_1 \times R_1 + c_2 \times R_2 + c_3 \times R_3$ 
27   end
28 end
29 denote the number of like for each comment as  $w_j$  ;
30  $L = \sum_j^n R_j^f \times (w_j + 1)$ 

```

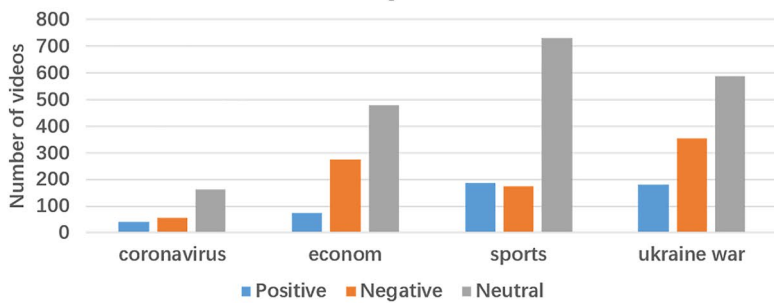
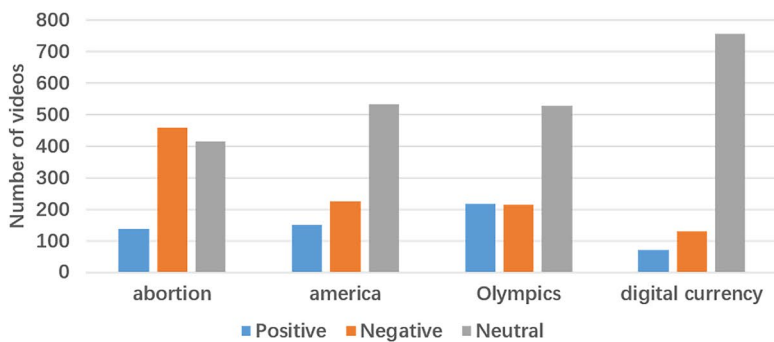
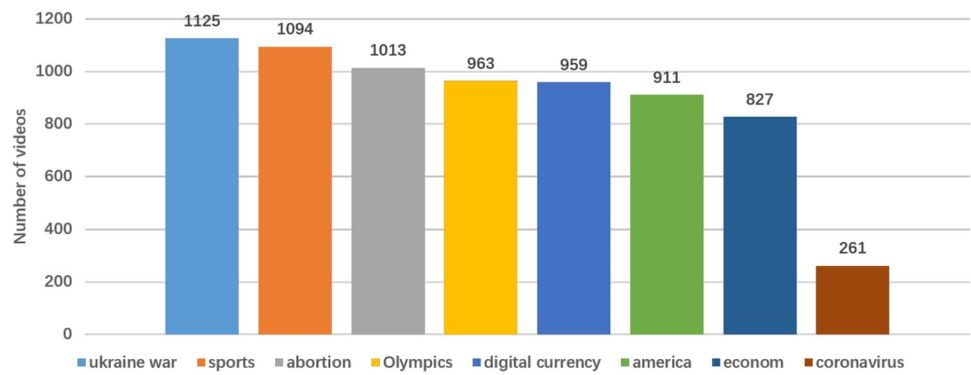
Once we obtain the inference results for all the comments, we require a strategy to aggregate them and derive the final public induced emotion. In this case, the number of likes associated with the comments becomes a crucial consideration, as it serves as an indicator of agreement from others. Therefore, we introduce a weighting mechanism by multiplying the number of likes (plus 1 to account for comments with zero likes) with the probability distribution vector of the comment. These weighted vectors are then accumulated together. Subsequently, we normalize the result, yielding a probability distribution indicative of the final public induced emotion on the micro video. To provide a clear representation of the label generation strategy, we present it in the form of pseudo-code, as demonstrated in Algorithm 1

3.4 Dataset statistics

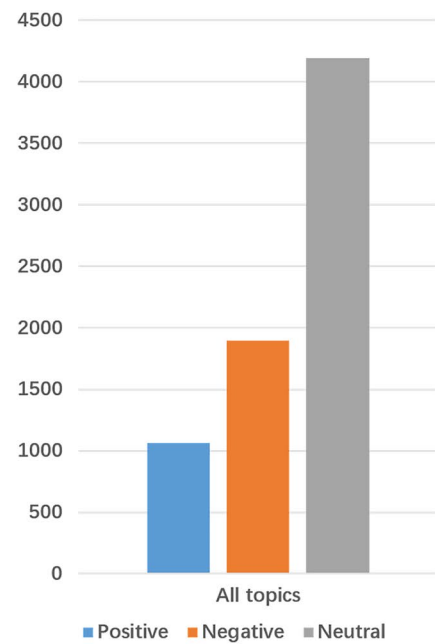
3.4.1 Dataset size

This extensive dataset, known as MVIndEmo, has undergone meticulous cleaning and filtering processes. We have gathered a substantial amount of data from the TikTok website. As a result, we have curated a selection of 7,153 videos that cover eight popular topics. Each video is accompanied by public-induced polarity labels and emotion classification labels. The distribution of video counts across these chosen topics can be observed in Fig. 3.

Fig. 3 The distribution of the number of videos on each topic



a) The distribution of polarity labels on each individual topic



b) The overall distribution of polarity labels

Fig. 4 The distribution of the polarity labels

3.4.2 Distribution of labels

The distribution of polar labels has been thoroughly analyzed and presented in Fig. 4. Subfigure (a) showcases the label distribution for each individual topic, while subfigure (b) provides insights into the overall label distribution. Upon inspection of Fig. 4, it becomes apparent that neutral sentiment prevails across most topics. Furthermore, negative sentiments tend to outnumber positive sentiments, which aligns with the real-world dynamics surrounding the selected topics. Likewise, the distribution of emotion classification labels is depicted in Fig. 5. Subfigure (b) reveals that joy and anger are the prevailing emotions observed within the dataset. Subfigure (a) emphasizes that only the Olympics, digital currency, and economy topics elicit a predominantly joyous mood among viewers, while the remaining topics tend to

provoke a more angered state. It is worth noting that this imbalanced distribution may be attributed to the nature of the chosen social topics. Micro videos centered around daily life are more likely to evoke emotions of love or surprise, while social topics tend to elicit stronger reactions of anger. In future work, we will expand the topics and hashtag to balance the label distribution, such as #pet, #baby and so on.

3.4.3 Licensing

The MVIndEmo dataset is in the public domain. We released the MVIndEmo data set under a CC BY-SA 4.0 license. The MVIndEmo data set is for research purposes only to promote research about the induced emotion analysis of micro video.

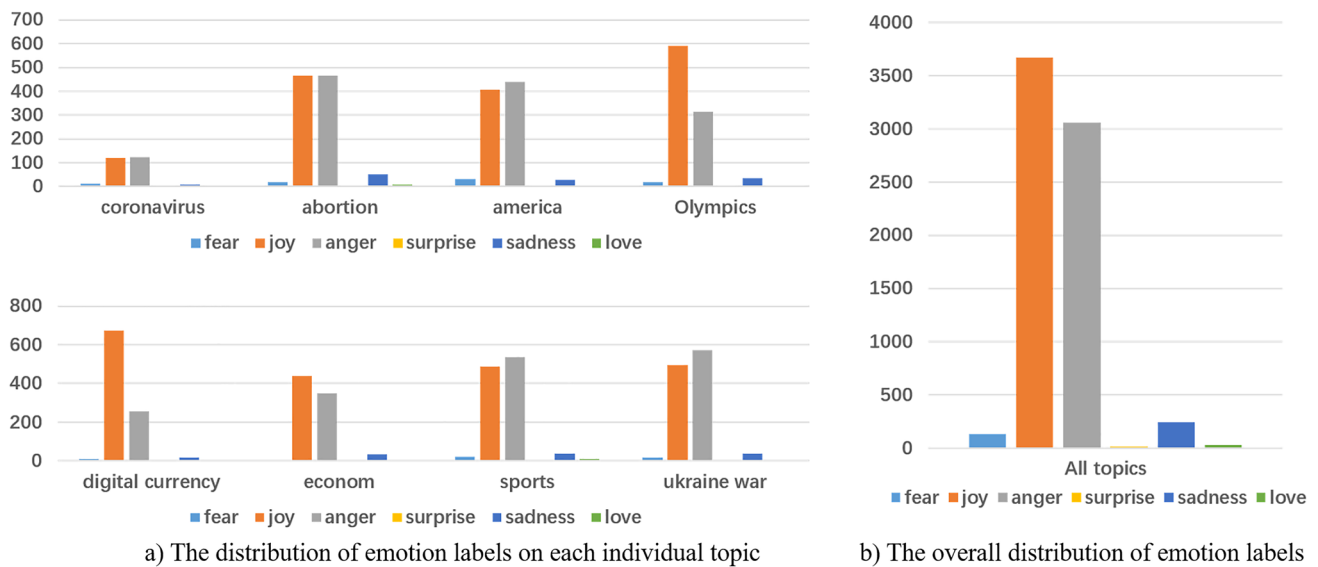


Fig. 5 The distribution of the emotion classification labels

3.4.4 Ethical considerations

This MVIndEmo dataset comes from the data collected on TikTok. In order to avoid access to personal privacy data, we only provide the webpage URL of the micro video and the public-induced emotion label that we generate according to the comments, but not the original video file and comment information. Researchers can visit Corresponding web pages to obtain the micro video data. The dataset is available on <https://github.com/inspur-hsslab/NeurIPS-Dataset-Induced-Emotion.git>.

4 Conclusion

In this research paper, we introduce the MVIndEmo dataset, a valuable resource for predicting public-induced emotions on micro videos of social media platforms. This dataset serves as the foundation for our innovative research task, which focuses on predicting the emotions induced by micro videos among the general public. To facilitate the prediction task, we have established two specific objectives: (1) emotion polarity prediction and (2) emotion classification. For each objective, we have provided two forms of labels: binary classification labels and soft labels that represent the probability distribution of induced emotions. We aim to create a more flexible benchmark for evaluating prediction models.

The dataset originates from TikTok, a popular social media platform known for its short, user-generated videos. Each micro video page on TikTok consists of the

video itself, a description accompanied by hashtags, comments from viewers, and likes on those comments. Users express their opinions about the micro video through comments, while the engagement level is measured by the number of likes received on those comments. Leveraging these comments, we have devised an automatic labeling strategy to generate the induced emotion labels for each micro video. The dataset encompasses a total of 7,153 micro videos, each enriched with corresponding labels. To offer a comprehensive overview, we have provided statistical analysis of the dataset, presenting the overall characteristics and composition. To the best of our knowledge, this dataset is the first of its kind for predicting public-induced emotions on micro videos. It stands as a significant contribution in the domain of sentiment analysis, particularly in the realm of induced emotions. By making this dataset publicly available, we hope to inspire new research avenues and support scholars who specialize in social media analysis and multi-modal information comprehension.

Acknowledgements Thanks for the support from the colleagues, especially Lu Liu, Yingjie Zhang and Ruidong Yan.

Author Contributions Conceptualization, QJ and YZ; methodology, QJ; software, XC, QJ and DW; validation, BF, and CX; formal analysis, YW and QJ; investigation, QJ; resources, YZ and ZG; data curation, DW and XC; writing-original draft preparation, QJ; writing-review and editing, BF; visualization, CX; supervision, RL; project administration, ZG and YW; funding acquisition, RL. All authors have read and agreed to the published version of the manuscript.

Funding Please add: This research was funded by the National Key Research and Development Program of China (No.2021ZD0113004)

grant funded by the Ministry of Science and Technology of the People's Republic of China.

Availability of data and materials The data could be available at “<https://github.com/inspur-hsslab/NeurIPS-Dataset-Induced-Emotion-gita>.”

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aljanaki, A., Wiering, F., Veltkamp, R.C.: Studying emotion induced by music through a crowdsourcing game. *Inf. Process. Manag.* **52**(1), 115–128 (2016)
- Song, Y., Dixon, S., Pearce, M.T., Halpern, A.R.: Perceived and induced emotion responses to popular music: categorical and dimensional models. *Music Percept.* **33**(4), 472–492 (2016)
- Baveye, Y., Bettinelli, J.-N., Dellandréa, E., Chen, L., Chamaret, C.: A large video database for computational models of induced emotion. In: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, IEEE, pp. 13–18 (2013)
- Teixeira, T., Wedel, M., Pieters, R.: Emotion-induced engagement in internet video advertisements. *J. Mark. Res.* **49**(2), 144–159 (2012)
- Sukhwai, P.C., Kankanalli, A.: Determining containment policy impacts on public sentiment during the pandemic using social media data. *Proc. Natl. Acad. Sci.* **119**(19), 2117292119 (2022)
- Peng, K.-C., Chen, T., Sadovnik, A., Gallagher, A.C.: A mixed bag of emotions: Model, predict, and transfer emotion distributions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 860–868 (2015)
- Zhao, S., Yao, H., Gao, Y., Ding, G., Chua, T.-S.: Predicting personalized image emotion perceptions in social networks. *IEEE Trans. Affect. Comput.* **9**(4), 526–540 (2016)
- Arasteh, S.T., Monajem, M., Christlein, V., Heinrich, P., Nicolaou, A., Boldaji, H.N., Lotfinia, M., Evert, S.: How will your tweet be received? predicting the sentiment polarity of tweet replies. In: 2021 IEEE 15th International Conference on Semantic Computing (ICSC), IEEE, pp. 370–373 (2021)
- Shmueli, B., Ray, S., Ku, L.-W.: Happy dance, slow clap: Using reaction gifs to predict induced affect on twitter. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), pp. 395–401 (2021)
- Nakov, P., Kozareva, Z., Ritter, A., Rosenthal, S., Stoyanov, V., Wilson, T.: SemEval-2013 task 2: Sentiment analysis in Twitter. In: *SEM 2013 - 2nd Joint Conference on Lexical and Computational Semantics 2(SemEval), pp. 312–320 (2013)
- Saravia, E., Liu, H.-C.T., Huang, Y.-H., Wu, J., Chen, Y.-S.: Carer: Contextualized affect representations for emotion recognition. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pp. 3687–3697 (2018)
- Kallinen, K., Ravaja, N.: Emotion perceived and emotion felt: same and different. *Music. Sci.* **10**(2), 191–213 (2006)
- Tian, L., Muszynski, M., Lai, C., Moore, J.D., Kostoulas, T., Lombardo, P., Pun, T., Chanel, G.: Recognizing induced emotions of movie audiences: are induced and perceived emotions the same? In: 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE, pp. 28–35 (2017)
- Alqahtani, G., Alothaim, A.: Predicting emotions in online social networks: challenges and opportunities. *Multimed. Tools Appl.* **81**(7), 9567–9605 (2022)
- Benini, S., Canini, L., Leonardi, R.: A connotative space for supporting movie affective recommendation. *IEEE Trans. Multimed.* **13**(6), 1356–1370 (2011)
- Muszynski, M., Tian, L., Lai, C., Moore, J.D., Kostoulas, T., Lombardo, P., Pun, T., Chanel, G.: Recognizing induced emotions of movie audiences from multimodal information. *IEEE Trans. Affect. Comput.* **12**(1), 36–52 (2019)
- Liu, Y.-J., Yu, M., Zhao, G., Song, J., Ge, Y., Shi, Y.: Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. Affect. Comput.* **9**(4), 550–562 (2017)
- Katz, P., Singleton, M., Wicentowski, R.: Swat-mp: the semeval-2007 systems for task 5 and task 14. In: Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), pp. 308–313 (2007)
- Strapparava, C., Mihalcea, R.: Semeval-2007 task 14: Affective text. In: Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), pp. 70–74 (2007)
- Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., Yu, Y.: Joint emotion-topic modeling for social affective text mining. In: 2009 Ninth IEEE International Conference on Data Mining, IEEE, pp. 699–704 (2009)
- Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., Yu, Y.: Mining social emotions from affective text. *IEEE Trans. Knowl. Data Eng.* **24**(9), 1658–1670 (2011)
- Rao, Y., Li, Q., Wenyan, L., Wu, Q., Quan, X.: Affective topic model for social emotion detection. *Neural Netw.* **58**, 29–37 (2014)
- Clos, J., Bandhakavi, A., Wiratunga, N., Cabanac, G.: Predicting emotional reaction in social networks. In: European Conference on Information Retrieval, pp. 527–533. Springer, Cham (2017)
- Hariri, R.H., Fredericks, E.M., Bowers, K.M.: Uncertainty in big data analytics: survey, opportunities, and challenges. *J. Big Data* **6**(1), 1–16 (2019)
- Brown, D.G.: Classification and boundary vagueness in mapping presettlement forest types. *Int. J. Geogr. Inf. Sci.* **12**(2), 105–129 (1998)
- Wang, Q.A.: Probability distribution and entropy as a measure of uncertainty. *J. Phys. A* **41**(6), 065004 (2008)
- Northcutt, C., Jiang, L., Chuang, I.: Confident learning: estimating uncertainty in dataset labels. *J. Artif. Intell. Res.* **70**, 1373–1411 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.