RESEARCH-ARTICLE

# SuicidEmoji: Derived Emoji Dataset and Tasks for Suicide-Related Social Content

**TIANLIN ZHANG**, The University of Manchester, Manchester, Greater Manchester, U.K.

**KAILAI YANG**, The University of Manchester, Manchester, Greater Manchester, U.K.

**SHAOXIONG JI**, University of Helsinki, Helsinki, Uusimaa, Finland

**BOYANG LIU**, The University of Manchester, Manchester, Greater Manchester, U.K.

**QIANQIAN XIE**, The University of Manchester, Manchester, Greater Manchester, U.K.

**SOPHIA ANANIADOU**, The University of Manchester, Manchester, Greater Manchester, U.K.



**Published:** 10 July 2024



SIGIR 2024: The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval
*July 14 - 18, 2024*
*Washington DC, USA*

**Conference Sponsors:**
SIGIR

SIGIR '24: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (July 2024)
https://doi.org/10.1145/3626772.3657852
ISBN: 9798400704314

# SuicidEmoji: Derived Emoji Dataset and Tasks for Suicide-Related Social Content

Tianlin Zhang*
The University of Manchester
Manchester, UK
tianlin.zhang@manchester.ac.uk

Kailai Yang
The University of Manchester
Manchester, UK
kailai.yang@manchester.ac.uk

Shaoxiong Ji
University of Helsinki
Helsinki, Finland
shaoxiong.ji@helsinki.fi

Boyang Liu
The University of Manchester
Manchester, UK
boyang.liu-2@manchester.ac.uk

Qianqian Xie*
The University of Manchester
Manchester, UK
xqq.sincere@gmail.com

Sophia Ananiadou
The University of Manchester
Manchester, UK
Archimedes/Athena RC
Greece
sophia.ananiadou@manchester.ac.uk

## ABSTRACT

Early suicidal ideation detection using social media is crucial for mental health surveillance. Simultaneously, emojis from the posts can help us better understand users' emotions and predict mental health conditions. However, research in emoji-based suicide analysis remains underexplored, with few resources available, which can restrict the development of studying emoji usage patterns among users with suicidal ideation. In this work, we build a derived suicide-related emoji dataset named SuicidEmoji, which contains 25k emoji posts (2,329 suicide-related posts and 22,722 posts for the control group users) filtered from about 1.3 million crawled Reddit data. To the best of our knowledge, SuicidEmoji is the first suicide-related emoji dataset. Based on SuicidEmoji, we propose two novel tasks: emoji-aware suicidal ideation detection and emoji prediction, for which we build two benchmark subdatasets from SuicidEmoji to evaluate the performance of advanced methods including pre-trained language models (PLMs) and large language models (LLMs). We analyze the experimental results of two PLMs and the highly capable LLMs, which reveal the significance and challenges of emoji-based suicide-related NLP tasks. The dataset is avaliable at https://github.com/TianlinZhang668/SuicidEmoji.

## CCS CONCEPTS

• **Information systems → Information retrieval**; **Retrieval tasks and goals**; **Clustering and classification**.

## KEYWORDS

Suicidal Ideation Detection; Emojis; Mental Health; Social Media

**ACM Reference Format:**

Tianlin Zhang, Kailai Yang, Shaoxiong Ji, Boyang Liu, Qianqian Xie, and Sophia Ananiadou. 2024. SuicidEmoji: Derived Emoji Dataset and Tasks for Suicide-Related Social Content. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24), July 14–18, 2024, Washington, DC, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3626772.3657852

*Corresponding author



*SIGIR '24, July 14–18, 2024, Washington, DC, USA*

ACM ISBN 979-8-4007-0431-4/24/07
https://doi.org/10.1145/3626772.3657852

## 1 INTRODUCTION

Suicide has become one of the most serious public health problems. According to the report from World Health Organization (WHO) [1], more than 700,000 individuals died by suicide every year globally, and suicide is the fourth leading cause of death among teenagers. Additionally, many people with suicidal ideation do not seek professional diagnosis and treatment due to a lack of awareness or stigma [27]. With the development of social media in today's communication, people are increasingly using social media platforms, such as Reddit, to share their feelings and even communicate suicidal tendencies. This makes social media posts a valuable resource for mental health surveillance and suicide prevention [13, 31]. Therefore, it is necessary to understand people's mental states and detect potential suicidal ideation early through their posted content.

Emojis are graphical symbols widely used in modern text communication [20]. Recently, the research on emojis has been receiving increasing attention, as emojis play important roles on social media and carry rich semantic and emotional information [3]. Meanwhile, existing studies indicate that emojis can help us understand the emotions expressed in the texts and provide useful clues to predict mental illness [8, 19, 21]. As shown in Figure 1, the emoji usage differs slightly across different subreddits (user-created communities of interest in Reddit), and the posts with potential suicidal tendencies are more likely to contain emoji with negative emotions (e.g., 💔, 😞) frequently. However, most suicide-related datasets remove emojis during data pre-processing [1, 30], which seriously ignores the aforementioned usage patterns among users with suicidal ideation.

To bridge the gap of suicide-related emoji resources and spur future emoji research, we build a derived dataset named SuicidEmoji, which contains 25,051 emoji posts (2,329 suicide-related posts and 22,722 posts from control group users) filtered from approximately 1.3 million crawled Reddit data. To the best of our knowledge, SuicidEmoji is the first suicide-related emoji dataset. In particular, to better understand the challenges in emoji-based natural language

[1] https://www.who.int/news-room/fact-sheets/detail/suicide

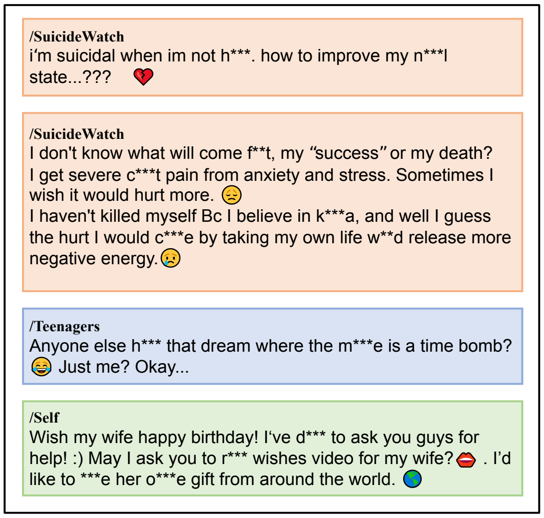
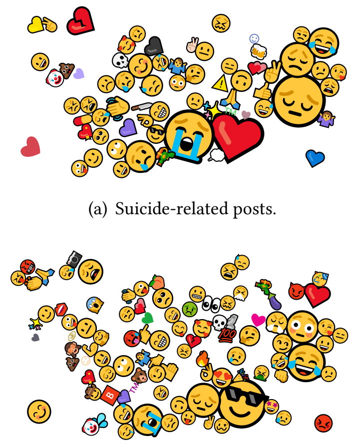



**Figure 1: Some examples with emojis posted by users in different subreddits (e.g., /SuicideWatch, /Teenagers, /Self).**



(a) Suicide-related posts.

(b) Control group users' posts.

**Figure 2: The emoji word clouds from suicide-related and control posts, displaying the most frequent emojis.**

processing (NLP) tasks, we propose two novel tasks: emoji-aware suicidal ideation detection and emoji prediction, based on two benchmark subdatasets from SuicidEmoji. Furthermore, we evaluate several state-of-the-art (SOTA) models, including pre-trained language models (PLMs) and large language model (LLM) such as ChatGPT [2], on the proposed benchmarks. The experimental results indicate that emoji features are beneficial for suicidal ideation detection. Additionally, the current SOTA methods do not perform quite well on the emoji prediction task, which illustrates that it still needs efforts, such as developing specific models to better understand the meaning behind emojis.

In summary, this paper makes the following contributions: 1) We release the first emoji-based suicide-related dataset, SuicidEmoji, containing 25k emoji posts, to facilitate the development of emoji-based NLP for mental health applications. 2) We propose two novel tasks: emoji-aware suicidal ideation detection and emoji prediction, for which we build two benchmark subdatasets from SuicidEmoji. 3) We evaluate the performance of several SOTA methods including PLMs and ChatGPT on the build two benchmarks. The experimental results reveal the significance and challenges of emoji-based suicide-related tasks.

# 2 THE SUICIDEMOJI DATASET

## 2.1 Data Collection

We obtain the source crawled Reddit data from two large-scale suicidal text corpora, including SuicideReddit [23] and Robin [9], both of which are collected from Reddit using pushshift API [5]. The SuicideReddit dataset collects suicide-related posts from the "/SuicideWatch" subreddit (Dec, 2008 ~Jan, 2021), where users with suicidal ideation often post a cry, and the control group users' posts are from the "/teenagers" subreddit. Similarly, the Robin dataset is also sourced from the "/SuicideWatch" subreddit and 13 other additional subreddits (e.g., /CasualConversation, /self, /TIFU) commencing in 2019. Although these suicidal ideation posts from "/SuicideWatch" do not have human annotation, the quality of self-labeling remains relatively reliable, given the large-scale nature of the corpus and the validation conducted in previous research [9].

After removing duplicate posts from both datasets and obtaining about 1.3 million posts, we only keep the posts that contain emojis. Then, we use *demoji* package[2] to decode them since most emojis are encoded in the Unicode standard. As a result, we collect 25,051 emoji instances containing 2,329 suicide-related posts and 22,722 posts from control group.

## 2.2 Data Analysis

The detailed statistics of our dataset are presented in Table 1. We can observe that the users with suicidal ideation tend to post longer messages (average 225.82 words) and use fewer emojis (average 2.60 per post) compared to control group users. Particularly, in the "/SuicideWatch" subreddit, the users often use negative emotion emojis, such as pensive face (😔), loudly crying face (😭) and broken heart (💔) to express their feelings, which is also illustrated in the emoji word clouds from Figure 2. According to the emoji sentiment scores from [17], control group users prefer to use positive emojis (e.g., smiling face with sunglasses (😎), face with tears of joy (😂)). Additionally, the frequently co-used emojis further reflect that a user with a potential suicide attempt uses negative emojis with a much higher probability than control group users. For example, post "*I *** my life is at the end of the l**. I can't do a***e.. so, I want to say goodbye*👋😔" using waving hand emoji and pensive face emoji to convey a profound sense of hopelessness. In conclusion, some emoji patterns are correlated with suicidal ideation.

[2]https://pypi.org/project/demoji/

| | Suicide | | | | | | | | | | Control | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **# of posts** | 2,329 | | | | | | | | | | 2,2722 | | | | | | | | | |
| **Avg. length** | 225.82 | | | | | | | | | | 123.94 | | | | | | | | | |
| **Avg. # of emojis** | 2.60 | | | | | | | | | | 10.14 | | | | | | | | | |
| **Avg. # of different emojis** | 1.37 | | | | | | | | | | 2.30 | | | | | | | | | |
| **Top 10 emojis** | [😭 | 🙂 | ❤️ | 😔 | 😂 | 😢 | 😞 | 💔 | 💩 | 😊 ] | [🤡 | 🗿 | 🦀 | 🎱 | 😂 | 🇺🇸 | 💰 | 😎 | 🅱️ | 😭 ] |
| **Top 10 emojis (rm dup.)** | [❤️ | 😔 | 😭 | 😞 | 😢 | 😂 | 💔 | 🖤 | ✌️ | 😊 ] | [😎 | 😂 | 😔 | 😳 | 😭 | ❤️ | 😩 | 😊 | 😤 | 🖤 ] |
| **Emoji sentiment scores** | 0.66 | -0.15 | -0.09 | -0.15 | 0.01 | 0.22 | -0.12 | 0.75 | 0.46 | 0.64 | 0.49 | 0.22 | -0.15 | 0.02 | -0.09 | 0.66 | -0.37 | 0.64 | null | 0.75 |
| **Top 5 co-used emojis** | [👋, 😔 | | 💔, 😭 | | 😢, 😭 | | 💔, 😔 | | 😔, 😢 ] | | [👈, 👉 | | 💦, 😩 | | 👉,😎 | | 😂, 🤣 | | 👁️, 👄 ] | |

**Table 1: Detailed statistics of the SuicidEmoji dataset, including the number (#) of posts, average (Avg.) length of posts, average number of emojis in each post, average number of different emojis in each post, top 10 emojis, top 10 emojis after removing duplicate, the corresponding emoji sentiment scores (range from -1 to 1 to express the polarity and intensity of sentiment) and top 5 co-used emojis.**

## 2.3 Evaluation Benchmark

To facilitate the development of emoji-based NLP for mental health applications, based on the SuicidEmoji dataset, we build a benchmark containing two novel tasks: emoji-aware suicidal ideation detection and emoji prediction.

**Suicidal Ideation Detection (SID)** This task detects whether the posts with emojis express potential suicidal ideation. Based on SuicidEmoji, we randomly select 10,000 raw data from the control set and combine them with 2,329 suicide-related posts to create a new dataset $D_{SID}$, thereby alleviating the impact of class imbalance.

**Emoji Prediction (EP)** Following related emoji research [4, 7], we build an emoji prediction task to better understand emoji usage patterns by different users and the feelings behind the text. The task aims to predict emoji considering the contextual content of one post; thus, we exclusively select posts with one emoji [4] from the SuicidEmoji dataset. The new dataset $D_{EP}$ consists of 886 suicide-related posts, 6,248 control posts, and an additional subset of 887 control posts matching the scale of the suicide set. We choose the 20 most frequent emojis from each set, and the emoji labels, along with their corresponding percentages, are shown in Table 2.

| | Suicide | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **# of posts** | 886 | | | | | | | | | |
| | ❤️ | 😔 | 😭 | 😞 | ✌️ | 😊 | 😢 | 😂 | 💔 | 🙃 |
| **Lables** | 17.4 | 14.0 | 10.0 | 8.6 | 5.8 | 5.1 | 4.8 | 4.7 | 4.3 | 3.1 |
| **(%)** | 👋 | 🖤 | 🙄 | 👍 | 😓 | 🤔 | 😕 | 😁 | 😅 | 😩 |
| | 3.1 | 2.6 | 2.4 | 2.4 | 2.2 | 2.2 | 2.0 | 2.0 | 1.7 | 1.6 |
| | **Control/sub_Control** | | | | | | | | | |
| **# of posts** | 6248/887 | | | | | | | | | |
| | 😎 | 😂 | 😔 | ❤️ | 😭 | 😳 | 😊 | 😐 | 😁 | 😅 |
| **Lables** | 16.5 | 11.0 | 10.2 | 7.1 | 7.0 | 6.7 | 6.0 | 3.5 | 3.4 | 3.4 |
| **(%)** | 🤔 | 🙄 | 🙃 | 🥺 | 😢 | 😩 | 😤 | 😏 | 👍 | 😀 |
| | 3.3 | 3.1 | 2.8 | 2.5 | 2.4 | 2.4 | 2.3 | 2.3 | 2.1 | 2.0 |

**Table 2: The 20 most frequent emojis with their corresponding percentages of each set.**

# 3 EXPERIMENTS

## 3.1 Baseline Models

For baselines, we evaluate two PLMs, including BERT [15] and mental healthcare-related MentalBERT [14], fine-tuned for our tasks. Meanwhile, due to the strong general language processing abilities exhibited by LLMs [16, 29], we also conduct performance tests on the representative models, Falcon (Falcon-7b-instruct) [24], LLaMA (LLaMA-13B-chat) [28], ChatGPT (gpt3.5-turbo) [3], using $D_{SID}$ and $D_{EP}$ datasets.

## 3.2 Settings

On the SID task, we have four different settings for the model inputs: *original*, *removed*, *emoji* and *emoji_text*. ***orginal***: original input - since most emojis are not in the vocabulary of BERT and MentalBERT, the tokenizer tokenizes emojis as [UNK] tokens. ChatGPT can recognize and process emojis. ***removed***: emojis removal - emojis are removed from the input text. ***emoji***: emojis added - emojis are added to the vocabulary with initial random values for the model's training. ***emoji_text***: emoji replacement - emojis are converted into textual descriptions using *emoji* package[4] and included in the input. For example, 👍 is converted into thumbs up. Specifically, for the evaluation of LLMs, we design the following prompt: *Post: "[Post]". Consider this post to answer the question: Does the post express the author's suicidal ideation? Only return Yes or No.* It is relatively more critical for a detection model to ensure that users with suicidal ideation are not misclassified [14]. Hence, we report the recall and weighted F1 scores for both the suicide and average categories.

On the EP task, we also evaluate BERT, MentalBERT and LLMs with the prompt: *Post: "[Post]". Consider this post to select one of the most suitable emoji from the following emoji set: [List]. Only return the assigned emoji.* The *[List]* are the predefined labels presented in Table 2. We leverage micro F1, macro F1 and weighted F1 scores as evaluation metrics.

We conduct PLMs' experiments on the Nvidia Tesla V100 GPU with 16GB of memory. We use grid search to explore the optimal hyper-parameters: the batch size is 16, and the learning rate is $5e^{-5}$. We fine-tune the models with AdamW optimizer for 20 epochs and employ an early stopping strategy. All experimental results are derived from the average scores obtained through a 10-fold cross-validation.

[3]https://openai.com/blog/chatgpt
[4]https://pypi.org/project/emoji/

### 3.3 Overall Results

The results of the SID task are presented in Table 3. We can observe that BERT_*emoji_name* and MentalBERT_*emoji* models have achieved the best performance. When emojis are removed or tokenized with [UNK] tokens, the predictive performance deteriorates because the models neglect emoji information. At the same time, BERT_*emoji_text* outperforms other models in terms of recall and F1 scores for the suicide category, which demonstrates the descriptive information conveyed by emojis is indeed beneficial for suicidal ideation detection and is similar to previous research findings [26]. On the average results, MentalBERT_*emoji* achieves the highest performance, further illustrating the significance of emojis and the meaningfulness of the constructed dataset. In addition, we find that in the zero-shot settings, ChatGPT_*original* also exhibits acceptable results in the SID task, approaching the performance of fine-tuned PLMs, which indicates ChatGPT's ability to recognize emojis and the good quality of our data [11].

| | Suicide | | Average | |
|---|---|---|---|---|
| **Model/Metric** | **R** | **F1** | **R** | **F1** |
| BERT_*original* | 78.3 | 84.4 | 92.8 | 92.6 |
| BERT_*removed* | 71.4 | 78.7 | 92.7 | 92.4 |
| BERT_*emoji* | 81.1 | 84.6 | 94.4 | 94.3 |
| BERT_*emoji_text* | **88.0** | **87.1** | **94.5** | **94.4** |
| MentalBERT_*original* | 81.2 | 81.6 | 93.5 | 93.5 |
| MentalBERT_*removed* | 79.4 | 82.4 | 93.6 | 93.5 |
| MentalBERT_*emoji* | **83.4** | **85.5** | **94.6** | **94.6** |
| MentalBERT_*emoji_text* | 80.0 | 84.1 | 94.3 | 94.2 |
| Falcon_*original* | 66.4 | 68.2 | 71.2 | 71.5 |
| Falcon_*removed* | 65.3 | 67.5 | 71.0 | 71.1 |
| Falcon_*emoji_text* | 66.5 | 69.3 | 72.5 | 72.8 |
| LLaMA_*original* | 70.3 | 75.5 | 87.6 | 87.5 |
| LLaMA_*removed* | 70.2 | 74.3 | 87.8 | 87.4 |
| LLaMA_*emoji_text* | 71.5 | 72.7 | 88.2 | 88.3 |
| ChatGPT_*original* | 72.6 | 80.3 | 93.2 | 93.0 |
| ChatGPT_*removed* | 70.6 | 79.4 | 93.1 | 92.7 |
| ChatGPT_*emoji_text* | 69.7 | 79.2 | 93.1 | 92.7 |

**Table 3: The test performances of baseline models on SID task. *original*: original input; *removed*: emojis removal; *emoji*: emojis added; *emoji_text*: replace with emoji descriptions. We highlight the top 2 results in bold.**

For the EP task, the results are shown in Table 4. We observe that all models achieve lower performance on the $D_{EP}$ dataset, revealing the challenges of emoji prediction tasks due to multiple interpretations of emojis and cultural variations [18, 25]. MentalBERT achieves remarkable results on the suicide set, which proves the effectiveness of pre-training in the mental health domain. Meanwhile, considering the performance differences between the suicide set and the sub_control set, we speculate that users with suicidal tendencies may have similar emoji usage patterns, reflecting the significance of exploring emoji features. Because of the increased number of training samples, the performance of the control set is better than that of the sub-control set. Furthermore, the assessment of results on the EP task also demonstrates that emoji prediction and user behavior understanding are still challenging for LLMs [10].

| **Dataset** | **Suicide** | | |
|---|---|---|---|
| **Model/Metric** | **MicroF1** | **MacroF1** | **WeightedF1** |
| BERT | 31.6 | 13.3 | 24.6 |
| MentalBERT | **36.8** | **26.4** | **35.5** |
| Falcon | 9.3 | 3.2 | 8.7 |
| LLaMA | 10.4 | 3.8 | 9.4 |
| ChatGPT | 12.0 | 4.0 | 9.6 |
| | **Control** | | |
| | **MicroF1** | **MacroF1** | **WeightedF1** |
| BERT | **31.1** | 12.8 | **24.1** |
| MentalBERT | 29.0 | 11.2 | 21.8 |
| Falcon | 14.3 | 9.8 | 12.3 |
| LLaMA | 16.9 | 11.0 | 15.3 |
| ChatGPT | 20.0 | **14.6** | 19.0 |
| | **sub_Control** | | |
| | **MicroF1** | **MacroF1** | **WeightedF1** |
| BERT | 26.1 | 7.2 | 16.2 |
| MentalBERT | **26.9** | 7.0 | 17.1 |
| Falcon | 12.5 | 8.2 | 10.5 |
| LLaMA | 14.1 | 8.6 | 10.9 |
| ChatGPT | 19.4 | **12.3** | **18.2** |

**Table 4: The test performances of baseline models on EP task. We highlight the optimal results in bold.**

Table 5 shows the confusion matrix of MentalBERT on the suicide test set intuitively. For top-5 labels, we can observe that the model achieves higher accuracy in predicting emoji (❤️) with larger data volumes. However, it frequently misclassifies emoji pensive face (😔) and emoji disappointed face (😞), primarily due to the high semantic similarity between these two emojis [12].

| | | Predicted label | | | | |
|---|---|---|---|---|---|---|
| | | ❤️ | 😔 | 😭 | 😞 | ✌️ |
| True label | ❤️ | 17 | 1 | 0 | 2 | 0 |
| | 😔 | 1 | 7 | 2 | 8 | 0 |
| | 😭 | 0 | 1 | 7 | 2 | 0 |
| | 😞 | 0 | 5 | 1 | 4 | 0 |
| | ✌️ | 3 | 2 | 0 | 0 | 2 |

**Table 5: The confusion matrix of MentalBERT on the suicide test set. We show the top-5 labels with the highest percentages.**

## 4 CONCLUSION

In this paper, we present the first suicide-related emoji dataset, SuicidEmoji, which facilitates in-depth analysis and future research on emoji-based suicide-related NLP tasks. Based on SuicidEmoji, we build two benchmark subdatasets to explore the challenges of emoji-aware suicidal ideation detection and emoji prediction tasks. The experimental results reveal the challenges in existing models in this research domain.

# 5 ETHICAL CONSIDERATION AND LIMITATIONS

To ensure that users' privacy is protected and anonymity is appropriately applied, we strictly follow the ethical principles and privacy protocols [22]. Meanwhile, to minimize misuse, we utilize the moderate disguising scheme [6] to paraphrase and obfuscate all examples in this paper. The datasets we used are publicly available and we also strictly follow the ethical principles and privacy protocols. In addition, The corpus we used was crawled in 2021 before the new Reddit Data API terms (Effective June 19, 2023. Last Revised April 18, 2023) released [5]. Notably, we recognize the profound sensitivity of using AI in detecting suicidal ideation and do not underestimate the potential consequences. Our research is motivated by studying emoji usage patterns among users with suicidal ideation, not to replace but to supplement existing mental health support systems, including peer support and non-traditional approaches acknowledged by the anti-psychiatry and critical psychiatry movements. In addition, the pre-trained models and final predicted results should only be used for non-clinical research, and the users who experience suicidal ideation should get assistance from mental health professionals.

In our study, several limitations should be acknowledged. First, the used datasets are from a single data source, Reddit. In addition, it is conceivable that the dataset may primarily reflect the situations in English-speaking countries, especially America (the stars and stripes emoji in Table 1), potentially leading to a limited perspective in terms of global representation and cultural diversity. Second, we do not account for some of the potential impacts, such as biases, while evaluating the performance of ChatGPT. We do not claim that clinicians will be replaced. We acknowledge the absence of direct involvement with mental health experts in the current phase of our work, we emphasize its importance and our imminent plan to integrate it in mental health applications in collaboration with experts. In future work, we will collect more high-quality emoji-based data from other social media platforms and measure ChatGPT's other safety criteria. Additionally, we will explore the role of SuicidEmoji in emoji disambiguation analysis.

## ACKNOWLEDGMENTS

This work is supported by the project JPNP20006 from New Energy and Industrial Technology Development Organization (NEDO), and this work has been partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGeneration EU Program..

## REFERENCES

[1] Theyazn HH Aldhyani, Saleh Nagi Alsubari, Ali Saleh Alshebami, Hasan Alkahtani, and Zeyad AT Ahmed. 2022. Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. *International journal of environmental research and public health* 19, 19 (2022), 12635.

[2] Florian Arendt, Benedikt Till, Martin Voracek, Stefanie Kirchner, Gernot Sonneck, Brigitte Naderer, Paul Pürcher, and Thomas Niederkrotenthaler. 2023. ChatGPT, Artificial Intelligence, and suicide prevention: A call for a targeted and concerted research effort. *Crisis: The Journal of Crisis Intervention and Suicide Prevention* 44, 5 (2023).

[3] Qiyu Bai, Qi Dan, Zhe Mu, and Maokun Yang. 2019. A systematic review of emoji: Current research and future perspectives. *Frontiers in psychology* 10 (2019), 2221.

[5] https://www.redditinc.com/policies/data-api-terms

[4] Francesco Barbieri, Jose Camacho-Collados, Francesco Ronzano, Luis Espinosa Anke, Miguel Ballesteros, Valerio Basile, Viviana Patti, and Horacio Saggion. 2018. Semeval 2018 task 2: Multilingual emoji prediction. In *Proceedings of the 12th international workshop on semantic evaluation*. 24–33.

[5] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. In *Proceedings of the international AAAI conference on web and social media*, Vol. 14. 830–839.

[6] Amy Bruckman. 2002. Studying the amateur artist: A perspective on disguising data collected in human subjects research on the Internet. *Ethics and Information Technology* 4 (2002), 217–231.

[7] Jose Camacho-Collados, Kiamehr Rezaee, Talayeh Riahi, Asahi Ushio, Daniel Loureiro, Dimosthenis Antypas, Joanne Boisson, Luis Espinosa Anke, Fangyu Liu, and Eugenio Martínez-Cámara. 2022. TweetNLP: Cutting-Edge Natural Language Processing for Social Media. In *Proceedings of the The 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. 38–49.

[8] Moumita Chatterjee, Piyush Kumar, Poulomi Samanta, and Dhrubasish Sarkar. 2022. Suicide ideation detection from online social media: A multi-modal feature based technique. *International Journal of Information Management Data Insights* 2, 2 (2022), 100103.

[9] Daniel DiPietro, Vivek Hazari, and Soroush Vosoughi. 2022. Robin: A Novel Online Suicidal Text Corpus of Substantial Breadth and Scale. *arXiv preprint arXiv:2209.05707* (2022).

[10] Hamideh Ghanadian, Isar Nejadgholi, and Hussein Al Osman. 2023. ChatGPT for Suicide Risk Assessment on Social Media: Quantitative Evaluation of Model Performance, Potentials and Limitations. *arXiv preprint arXiv:2306.09390* (2023).

[11] Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. 2023. Chatgpt outperforms crowd-workers for text-annotation tasks. *arXiv preprint arXiv:2303.15056* (2023).

[12] Sara R Jaeger, Christina M Roigard, David Jin, Leticia Vidal, and Gastón Ares. 2019. Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. *Food research international* 119 (2019), 895–907.

[13] Shaoxiong Ji, Shirui Pan, Xue Li, Erik Cambria, Guodong Long, and Zi Huang. 2020. Suicidal ideation detection: A review of machine learning methods and applications. *IEEE Transactions on Computational Social Systems* 8, 1 (2020), 214–226.

[14] Shaoxiong Ji, Tianlin Zhang, Luna Ansari, Jie Fu, Prayag Tiwari, and Erik Cambria. 2022. MentalBERT: Publicly Available Pretrained Language Models for Mental Healthcare. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. 7184–7190.

[15] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL-HLT*. 4171–4186.

[16] Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large language models are zero-shot reasoners. *Advances in neural information processing systems* 35 (2022), 22199–22213.

[17] Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. 2015. Sentiment of emojis. *PloS one* 10, 12 (2015), e0144296.

[18] Weicheng Ma, Ruibo Liu, Lili Wang, and Soroush Vosoughi. 2020. Emoji prediction: Extensions and benchmarking. *arXiv preprint arXiv:2007.07389* (2020).

[19] Davide Marengo, Michele Settanni, and Fabrizia Giannotta. 2019. Development and preliminary validation of an image-based instrument to assess depressive symptoms. *Psychiatry research* 279 (2019), 180–185.

[20] Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. "Blissfully happy" or "ready tofight": Varying interpretations of emoji. In *Proceedings of the international AAAI conference on web and social media*, Vol. 10. 259–268.

[21] Usman Naseem, Jinman Kim, Matloob Khushi, and Adam Dunn. 2023. Graph-Based Hierarchical Attention Network for Suicide Risk Detection on Social Media. In *Companion Proceedings of the ACM Web Conference 2023*. 995–1003.

[22] Jennifer Nicholas, Sandersan Onie, and Mark E Larsen. 2020. Ethics and privacy in social media research for mental health. *Current psychiatry reports* 22 (2020), 1–7.

[23] K Nikhileswar, D Vishal, L Sphoorthi, and S Fathimabi. 2021. Suicide ideation detection in social media forums. In *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE, 1741–1747.

[24] Guilherme Penedo, Quentin Malartic, Daniel Hesslow, Ruxandra Cojocaru, Alessandro Cappelli, Hamza Alobeidli, Baptiste Pannier, Ebtesam Almazrouei, and Julien Launay. 2023. The RefinedWeb dataset for Falcon LLM: outperforming curated corpora with web data, and web data only. *arXiv preprint arXiv:2306.01116* (2023).

[25] Matthew Shardlow, Luciano Gerber, and Raheel Nawaz. 2022. One emoji, many meanings: A corpus for the prediction and disambiguation of emoji sense. *Expert Systems with Applications* 198 (2022), 116862.

[26] Abhishek Singh, Eduardo Blanco, and Wei Jin. 2019. Incorporating emoji descriptions improves tweet classification. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 2096–2101.

[27] Samantha Tang, Natalie M Reily, Andrew F Arena, Philip J Batterham, Alison L Calear, Gregory L Carter, Andrew J Mackinnon, and Helen Christensen. 2022. People who die by suicide without receiving mental health services: a systematic review. *Frontiers in public health* 9 (2022), 736948.

[28] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).

[29] Kailai Yang, Shaoxiong Ji, Tianlin Zhang, Qianqian Xie, and Sophia Ananiadou. 2023. On the evaluations of chatgpt and emotion-enhanced prompting for mental health analysis. *arXiv preprint arXiv:2304.03347* (2023).

[30] Tianlin Zhang, Annika M Schoene, and Sophia Ananiadou. 2021. Automatic identification of suicide notes with a transformer-based deep learning model. *Internet interventions* 25 (2021), 100422.

[31] Tianlin Zhang, Annika M Schoene, Shaoxiong Ji, and Sophia Ananiadou. 2022. Natural language processing applied to mental illness detection: a narrative review. *NPJ digital medicine* 5, 1 (2022), 46.