

I'm out of breath from laughing! I think?

A dataset of COVID-19 Humor and its toxic variants

This is the pre-print version of the paper. The final copy will be available at <https://dl.acm.org/conference/www/proceedings>.

Neha Reddy Bogireddy^{*†}

George Mason University
Fairfax, VA, USA
nbogired@gmu.edu

Smriti Suresh^{*†}

Boston University
Boston, MA, USA
smritis@bu.edu

Sunny Rai[†]

University of Pennsylvania
Philadelphia, PA, USA
sunnyrai@seas.upenn.edu

ABSTRACT

Humor is a cognitive construct that predominantly evokes the feeling of mirth. During the COVID-19 pandemic, the situations that arouse out of the pandemic were so *incongruous* to the world we knew that even factual statements often had a humorous reaction. In this paper, we present a dataset of 2510 samples hand-annotated with labels such as humor *style*, *type*, *theme*, *target* and stereotypes formed or exploited while creating the humor in addition to 909 memes. Our dataset comprises Reddit posts, comments, Onion news headlines, real news headlines, and tweets. We evaluate the task of humor detection and maladaptive humor detection on state-of-the-art models namely RoBERTa and GPT-3. The finetuned models trained on our dataset show significant gains over zero-shot models including GPT-3 when detecting humor. Even though GPT-3 is good at generating meaningful explanations, we observed that it fails to detect maladaptive humor due to the absence of overt targets and profanities. We believe that the presented dataset will be helpful in designing computational methods for topical humor processing as it provides a unique sample set to study the theory of incongruity in a post-pandemic world. **The data is available at https://github.com/smritae01/Covid19_Humor.**

CCS CONCEPTS

• Computing methodologies → Language resources.

KEYWORDS

Humor Detection, Maladaptive Humor, Creative Text Processing, Topical Humor, Memes, Hate Speech, COVID-19

ACM Reference Format:

Neha Reddy Bogireddy, Smriti Suresh, and Sunny Rai. 2023. I'm out of breath from laughing! I think? A dataset of COVID-19 Humor and its toxic variants. In *Companion Proceedings of the ACM Web Conference 2023 (WWW '23 Companion)*, April 30-May 4, 2023, Austin, TX, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3543873.3587642>

^{*}Both authors contributed equally to this research.

[†]This work was performed at École Centrale School of Engineering, Mahindra University, Hyderabad, India.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WWW '23 Companion, April 30-May 4, 2023, Austin, TX, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9419-2/23/04...\$15.00

<https://doi.org/10.1145/3543873.3587642>

1 INTRODUCTION

Topical humor is defined as jokes '*framed around a specific topic related to current events or dealing with issues that are important or popular at the current time*' [36]. An example of topical humor is, 'An Indian family went into self-quarantine after eating lunch at their English friend's house as they couldn't taste anything.' Here, one needs to be aware of the symptom of 'loss of taste' when one's infected with COVID-19 along with the cultural details that Indians prefer relatively spicy food to understand the humor. Another example of a factual statement is 'Albany funeral home offers drive-thru services in wake of COVID-19 pandemic'. Here, the notion of *drive-thru* at funeral homes sounds *incongruent* as the funeral is about grieving together [9]. However, such was the situation during COVID-19 when people avoided the dead bodies of their loved ones for fear of infection. It thus becomes complex to determine incongruity as per the evolving world. Topical humor including the jokes made during COVID-19 demands temporal knowledge of the event to figure out the humorous sense of the text.

Topical humor offers a focused picture of how an event is altering views of human society. For instance, "Deprived of work by COVID, prison escapee turns himself in after 30 years of beach life on the lam". This text becomes humorous only when you realize the absurdity of returning to prison and the harsh reality of the lack of jobs during the COVID-19. Through this humor, one can understand emerging worldviews, evolving beliefs, and changing dynamics of socio-cultural conventions. When combined with the right style of argument and addressed to the right audience, topical humor acts as a tool for information sharing [27, 37]. Consider the piece, "Covid loves two things. 1: A dense population. 2: A dense population." mocking lack of COVID-19 appropriate behavior. Researchers have extensively studied the role of humor in diffusing stress [15, 19], encouraging solidarity [3] and information sharing [12] during COVID-19. The cultural load in these humorous pieces is often a challenge when identifying *good* humor vs *bad* humor. In fact, COVID-19 humor was widely used to target certain ethnic communities and groups (e.g. vaccinated vs anti-vaxxers) using expressions that were never seen before [22, 28].

In this paper, we present a dataset¹ of COVID-19 humor annotated with labels namely humor style (such as *pun*, *sarcasm* etc.), humor type (i.e. adaptive vs maladaptive humor), the theme (such as *illness*, *family* etc.), the target of the humor (such as *women*, *elderly* etc.) and stereotypes. We evaluate the performance of state of the art models namely RoBERTa and GPT-3 model on our dataset for two tasks namely humor detection and maladaptive humor detection. We believe that this dataset can serve multiple research

¹link will be added upon acceptance

Table 1: An overview of collected data

Source	#posts	Example
<i>r/Jokes</i>	806	The World Health Organization has declared that dogs cannot transmit Coronavirus, and there is no reason to quarantine dogs anymore. W.H.O. let the dogs out.
<i>r/Onion</i>	103	Parents Can't Tell If Pandemic Inhibited Toddler's Social Skills Or If He's Just Taking After Dad
<i>r/NottheOnion</i>	543	Convict refuses parole, says jail safer than outside in pandemic
Reddit Comments	130	I'm out of breath from laughing!I think
Twitter	15	My son hasn't been able to go out since he was 13 because of the pandemic. He's a quaran-teen.
Total (Textual Humor)	1597	
<i>r/CoronavirusMemes</i>	909	
Total (Textual + Memes)	2506	
Non Humorous	1563	Facing a COVID Crisis, Nepal Cries Out for Help
Total Samples	4069	

purposes. Firstly, it will be a useful resource to understand non-traditional, topical humor arising out of situations of significance around us. Secondly, the collected dataset is tagged for targets and themes which can help in understanding the signals of microaggression against a community or a group. Thirdly, the collected dataset underlines the relevance of evolving cultural information and connotations when dealing with topical humor. We discuss these points in detail in Section 5.

2 RELATED WORK

Humor is a crucial ingredient when resolving feelings of anxiety [2] and improving overall well being of human society [30]. Hussein and Aljamili [15] found humorous content including caricatures and memes mocking the pandemic and unusual situations arising out of it, helpful in relieving stress. Olah and Ford [27] discovered that individuals with self-enhancing humor style often engaged in protective behaviors such as *wearing a mask*, *washing hands* during COVID-19. Moreover, the meaning of 'good' COVID-19 humor varies with demographic factors and personality traits [4]. The growing use of dark themes including *death* and *elderly people* leads to a high aversion towards COVID-19 humor when compared to the usual humor. In fact, Vurayai [35] pointed out the increasing number of disparaging COVID-19 jokes targeting women.

Bruns and Burgess [5] emphasized the lack of research on the role of humor in information sharing regarding a crisis. Kuipers [17] named this specific type of humor as "disaster humor" as it appears soon after incidents and generally follows the media coverage of the facts. Dynel [12] analysed a set of 174 COVID-19 mask memes to track the evolving voices in the highly shared content on social media. Torres et al. [34] collected 107 COVID-19 humor scripts each for English and Tagalog. These scripts were coded with the humor type (e.g. *self-deprecating humor*, *malicious or sexual humor*, *aggressive etc.*), humor targets (e.g. *people*, *countries*, *biblical characters etc.*), humor subjects (e.g. *appearance*, *food*, *country*, *date etc.*) and humor style (e.g. *conversational*, *knock-knock*, *lyrical*, *narrative*, *one-liner*, *question and answer*). However, their analysis is qualitative, and the size of the sample is very small.

Even though there is an interesting collection of work on COVID-19 humor highlighting its varied purposes, the analysis is performed on a small sample. Through this work, we provide a rich, well-curated dataset for studying COVID-19 humor in English.

3 HUMOR DURING COVID-19

3.1 Data Collection

The subreddits *r/Jokes*, *r/CoronavirusMemes*, *r/Onion* headlines and *r/NottheOnion* headlines were chosen to extract humorous posts. In order to extract only COVID-19 related jokes, keywords namely *pandemic*, *covid19*, *coronavirus*, *lockdown*, *quarantine*, *covid*, *corona*, *SARS* were used within the subreddit thread. We observed that the comments under humorous posts often tend to be humorous themselves. Therefore, we also extracted these comments. PRAW (Python Reddit API Wrapper) was used to extract the posts from Reddit. A total of 67 humorous Twitter accounts [23] were queried using the above mentioned COVID-19 keywords to extract COVID-19 humorous tweets.

In addition, a total of 1321 non-humorous yet COVID-19 related posts, were collected from subreddits namely *r/Coronavirus*, *r/COVID19*, *r/Chinaflu*, *r/Pandemic*. These posts serve as negative samples for machine learning tasks. The overview of the posts collected from different subreddit threads is provided in Table 1. All retrieved posts (see Table 1) were retained for further analysis. Kindly note that we collected 909 memes on COVID-19 however, the annotation was only performed for textual humor which will be discussed in the rest of the paper.

3.1.1 Data Reliability. A text is considered humorous if the intent of the text is to be humorous. Since we only considered Reddit threads known for their humorous content, it is safe to assume that the collected data is humorous. Internet memes are often a humorous piece of culture [8] and we thus assume all collected memes to be humorous. To analyze the quality of the collected textual data, we asked a team of human annotators to manually tag

Table 2: Example: Annotation Labels

Text: Due to lack of protective measures during the pandemic, retail cashiers and bank tellers are refusing to go to work . It's a counter strike
Humor style: Pun
Humor Type: Adaptive
Theme: Pandemic
Target: Working Class
Stereotype: Not Applicable

Text: The Pandemic Has Closed Public Restrooms, and Many Have Nowhere to Go
Humor style: Pun
Humor Type: Maladaptive
Theme: Pandemic
Target: Poor
Stereotype: Not Applicable

randomly selected 30 posts from all threads namely *r/Jokes*, *r/Onion*, *r/NottheOnion* and the comments.

Critchley [7] defined humor as jokes (spoken or written words) and actions that elicit laughter or generate mirth. Human annotators were asked to independently determine if a given post is humorous or not while keeping this definition in mind and the intent of the author. The inter-annotator agreement (Krippendorff's alpha) was found to be 0.932 for *r/Jokes*, 0.512 for *r/Onion*, 0.20 for *r/NottheOnion*, and 0.71 for *comments*.

Humor is highly subjective and is often not appreciated by all [23]. Keeping this in mind, the thread *r/Jokes* has an almost perfect agreement. The comments have the agreement of 0.71 which is reasonably good. The intent of *r/Onion* headlines is to be funny but it is presented as news. We thus feel that annotators struggled to understand covert humor. Consider the onion headline: "Trump Blasts Dr. Fauci Over Repeated Negative Remarks About Coronavirus." This was marked as non-humorous by the annotators. However, this headline is clearly a sarcastic remark.

As expected, the thread *r/NottheOnion* has the lowest agreement. These headlines are real news however, the unusual human behavior during COVID-19 ascribed a humorous angle to these headlines. Please note, that these headlines were picked from Reddit threads discussing humorous situations. So these headlines were perceived as funny by few if not all. Nevertheless, we asked human annotators to annotate all posts in *r/NottheOnion* thread to ensure the quality of the dataset. 305 posts that were marked as non-humorous were filtered out from the *humorous* set. So a total of **1189 unique** humorous posts were considered for further analysis.

3.2 Data Annotation

The online social discourse during the COVID-19 pandemic provided a unique opportunity to analyze the evolving themes, targets, and inherent stereotypes in COVID-19 humor. We have a total of 1189 unique humorous posts out of 2510 posts which were annotated with the following labels:

- Humor Style,
- Humor Type that is, Maladaptive or Adaptive humor,
- Theme of humor,

Table 3: Top-5 labels for Adaptive and Maladaptive Humor

Adaptive Humor	Maladaptive Humor
<i>Humor Style (Top 5)</i>	
Pun (29.05%)	Dark (23.48%)
Exaggeration (13.15%)	Pun (10.98%)
Matter of fact (12.86%)	Irony (10.98%)
Irony (11.19%)	Matter of fact (9.47%)
Relief (8.73%)	Exaggeration (8.33%)
<i>Humor Theme (Top-5)</i>	
Pandemic (37.03%)	Pandemic (44.11%)
Illness (9.92%)	Illness (18.63%)
Life Change (9.33%)	Love/Sexual (8.36%)
Political and WHO (8.94%)	Life Change (7.22%)
Love/Sexual (6.38%)	Physical Appearance (5.32%)
<i>Humor Target (Top-5)</i>	
General (52.01%)	General (32.95%)
Authorities (11.58%)	Authorities (12.88%)
Anti-vaxxers (5.79%)	Infected (10.23%)
Working Class (4.32%)	Anti-vaxxers (6.82%)
Infected (3.83%)	Women (6.44%)

- Target of humor,
- Stereotypes if any.

Please see Appendix A for annotation guidelines and definitions. The authors identified a total of 14 themes and 17 targets after analyzing a random set of humorous samples. We discovered themes (such as *groceries*, *illness*, *hospital*) and targets (such as *china*, *infected*, *elderly*) unique to the COVID-19 period that wouldn't have been observed otherwise. If no clear target could be identified in a post, we asked the annotators to mark it as *general*.

3.2.1 Annotated Dataset. A total of 243 (20.6%) posts were marked as maladaptive humor out of 1189 humorous posts. All posts were assigned a humor style and a theme. A target was identified for 670 humorous posts out of 1189 posts. The remaining posts had no clear target and was there marked as general. The complete distribution of key themes and targets for humorous samples is provided in Table 4. We provide the top-5 humor styles, themes, and targets in Table 3.

Pun, *Exaggeration* and *Matter of fact* were the most common humor styles for adaptive humor whereas *Dark humor*, *Puns* and *irony* are the dominant humor styles for maladaptive humor. *Pandemic* and *Illness* were the top-2 themes for adaptive and maladaptive humor. However, maladaptive humor themes were more personal such as *Love/Sexual*, *Life Change*, and *Physical Appearance* when compared with the adaptive humor. *Infected* and *Anti-masks/vaxxers* were the more targeted groups for maladaptive humor in comparison to adaptive humor.

Stereotypes concerning certain targets and categories were also investigated in order to understand the prejudices and notions emerged during the pandemic. We identified 69 stereotypes from a sample of 700 posts. Race (21.7%), Chinese population (20.3%) and Nationality (20.3%) were the top-3 categories for stereotypes. See Table A4 for more details. It may be noted that identified stereotypes overlapped with annotated *targets* and *categories* such as Gender,

Table 4: Distribution for Theme and Target

Theme (#posts)	<i>Example</i>	Target (#posts)	<i>Example</i>
Pandemic (493)	Coronavirus ruining your plans for 2020? Save them for 2022!. Cause 2022 is 2020 too.	General/None (617)	Why can't ants get coronavirus. Because they have little anty-bodies
Illness (150)	In a year, we're all going to remember this pandemic and laugh about it... ..at least the ones left...	Authority (152)	Trump wasn't lying about the coronavirus disappearing in April. He simply forgot to say 2021.
Life Changes (114)	It all started with a bat. Then toilet paper. Now we're going nuts in quarantine.. We really have gone bat, shit, crazy.	Anti-maskers/vaxxers (77)	It's primarily killing the unvaccinated. The vast majority of which are idiots.
Political & WHO (104)	Coronavirus isn't Trumps Fault. Ebola wasn't Obama's Fault. Sars wasn't Brush's fault..... And only a handful of cases of herpes were Clinton's fault.	Infected (66)	Don't worry, there are experts who are trained to deal with the coronavirus.. We call them coroners.
Love/sex (87)	I really hope coronavirus can't spread through sex. It would be so lonely being the last man on Earth.	Working Class (50)	Coronavirus. I bet the guy who invented hand sanitizer is rubbing his hands together now.
Pop culture (56)	Why did Princess Leia contract coronavirus?. Because she went to woo Han.	Couples (47)	This quarantine thing is not working out.. My wife and I are becoming friends and I almost told her about my girlfriend.
Boredom (56)	I finished 5 books during the quarantine. That's a lot of coloring.	Celebrities (38)	what do you call an artist during a global pandemic? . Vincent Van Cough
Groceries (48)	I went for an interview, they asked me where do I see myself in five years. Apparently, still using the toilet rolls I stockpiled during the Coronavirus outbreak is not the right answer	China (35)	What's the difference between kung pao and the coronavirus?. One's Chinese take out, the other takes out Chinese.
Physical Appearance (49)	Covid lockdown has me gaining weight so I started a new diet. Its called Two weeks to flatten Your curves	Women (34)	Japanese mayor says men should grocery shop during pandemic as women 'take a longer time
Hospital & Doctor (40)	"Raincoat" stolen by drunk man from hospital turns out to be PPE kit; tests COVID-19 positive	Children (28)	This pandemic has given me a lot of time to get to know my kids better..or worse..
Work/Jobs (30)	My boss says he can't afford to pay me during this pandemic... . But luckily he says I can work for exposure.	Criminals (27)	2020 is a weird year. 5 years ago i entered a bank in a mask and i got forced to quarantine for years, today i got praised for it.. weird times to be alive
Media (30)	Fox News Limits Pandemic Coverage To Avoid Giving Coronavirus Notoriety It Craves	Family (24)	What's the difference between Coronavirus and my dad?. Coronavirus only takes five days to appear.
Education (16)	OHSAA coronavirus rules: Students can wrestle, but can't shake hands	Religion (22)	Why are Muslim youth the best at dealing with pandemics? . Because as a Quran-Teen, you always keep Allah by your self.
Work from home (10)	When the lockdown started, all I did was masturbate and watch TV all day.. After 3 weeks it got awkward and my coworkers decided to tell me how to turn my camera off on Zoom.	Men (20)	There's a Pandemic Sperm Shortage, and Sperm King. Megadonors Are in High Demand
		Health Workers (16)	Thankful for all the health care workers during this pandemic. Without dem, it would be a panic.
		Fans (12)	Why did everyone have Covid-19 at the KPop concert?. Because a symptom of Coronavirus is lack of taste.
		Poor (9)	This is the first year that we didn't go to Hawaii because of coronavirus.... Every other year we don't go because we can't afford it.
		Elderly (9)	Why did the Coronavirus cross the road?. Cuz it likes to walk old people to the other side.

Table 5: Performance Evaluation. Here, A(ccuracy), P(recision), R(ecall) and F1 score have the standard interpretation. * denotes Finetuned RoBERTa

Model	Task	A	P	R	F1
RoBERTa	Task-A	0.64	1	0.4	0.571
	Task-B	0.33	0.5	0.373	0.427
RoBERTa*	Task-A	0.92	0.95	0.92	0.93
	Task-B	0.67	0.5	0.24	0.33
GPT-3	Task-A	0.70	0.894	0.566	0.693
	Task-B	0.64	0.428	0.272	0.333

Race, Sexuality, Social-class, (Dis)ability, Age, Politics, Religion, Occupation, and the Chinese population. We discuss this in more detail in Section 5.

4 BASELINE MODELS

We evaluate the predictive performance of widely known neural language models on two tasks that are, A) *humor detection* and B) *maladaptive humor detection*. The test set for task A was formed by randomly sampling a set of 100 posts comprising humorous and non-humorous posts. For task B, we created a random sample of 100 humorous posts comprising adaptive and maladaptive samples. For Task-A, test set comprise 60 Humor and 40 Non Humor samples. For Task-B, test set has 67 adaptive humor and 33 maladaptive humor samples. The test set is common for all baseline models provided below:

- **RoBERTa (Zero-shot):** Using RoBERTa trained on Multi-Genre Natural Language Inference corpus, we specify the 'Sequence to classify' and class labels as illustrated here².
E.g. For Task-A: sequence_to_classify = "After this Covid Pandemic I told myself I really need to stop drinking... But I don't take advice from a drunk." and candidate_labels = ['humorous', 'non-humorous'].
- **Finetuned RoBERTa:** For Task-A, we finetune RoBERTa model on our dataset excluding the test set to identify humorous samples from non-humorous samples. For Task B, to combat the issue of having imbalanced dataset, we used a data augmentation technique called "back translation" for over-sampling the minority class (maladaptive) resulting in 478 maladaptive and 946 adaptive samples. We then finetune RoBERTa on these humorous samples to identify adaptive humor from maladaptive humor. For both tasks, we performed stratified sampling and train:val is 80 : 20. The dropout rate is 0.3, the learning rate is $2 * 10^{-5}$ with AdamW optimizer. The model was trained for 5 epochs.
- **GPT-3 (Zero-Shot):** Our objective is to evaluate the performance of simple prompts. The prompt for the task of humor detection (Task A) and maladaptive humor detection (Task B) are provided in Table 6 and Table 7.
The prompt is followed by a post in the test set. We set the engine to *text - davinci - 002*, Max tokens as 400, Temperature as 0, Top p as 1.0, Frequency penalty as 0.5, and Presence Penalty as 0.0.

²<https://huggingface.co/roberta-large-mnli>

Table 6: Responses by GPT-3 for Task-A

Prompt: Given the context of covid-19 pandemic, explain how a post is humorous or not.

(1) **China flight attendants advised to wear diapers for Covid protection** – [Label: humorous]

Response: *This post is not humorous, as it is a serious subject related to the covid-19 pandemic. The risk of infection is a real concern for flight attendants, and the suggestion of wearing diapers could be seen as an extreme measure to protect themselves. This post is highlighting the seriousness of the pandemic and the risks associated with it and is not intended to be humorous.*

(2) **Whats the difference between the coronavirus and 737MAX?. The coronavirus is airborne.** –[Label: humorous]

Response: *This post is not humorous as it makes a comparison between the serious global health crisis caused by the coronavirus and the recent Boeing 737 MAX plane crashes. The comparison trivializes the severity of the coronavirus pandemic, making light of the situation, which is not humorous.*

(3) **UNC system says no refunds on tuition, fees if pandemic forces return to online-only classes** – [Label: Non-humorous]

Response: *This post could be seen as humorous since it is making light of a serious situation. It implies that the University of North Carolina system has chosen not to refund tuition and fees in the event that students are forced to take classes online due to the pandemic. This could be seen as a joke since many students are already learning online due to the pandemic and could be seen as an ironic situation.*

4.1 Results

As humor is highly subjective, it is thus entirely not possible to have an objective definition of correctness when performing the evaluation. In this study, we consider a prediction to be correct if it matches the label assigned by human annotators. The performance evaluation for both tasks is provided in Table 5.

For Task-A, there is a significant improvement in the performance of finetuned RoBERTa for the humor detection model over zero-shot RoBERTa and GPT-3. RoBERTa model finetuned on our dataset has F1 score of 0.93, an improvement of 36% over the zero-shot RoBERTa model and 24% over GPT-3 zero-shot model.

For Task-B, RoBERTa zero shot has the best performance of 0.427 followed by GPT-3. We specifically found GPT-3 failing for the cases without explicit offensive words or targets. The finetuned RoBERTa marked all samples as adaptive and therefore had zero F1 score. This is likely due to the class imbalance between adaptive and maladaptive samples. In the present study, we limit the scope of our evaluation to cross-entropy loss and do not explore loss functions designed to handle class imbalance.

For GPT-3, the prompts were designed to also generate explanations. We provide a few such explanations in Table 6 and 7. The generated explanations were meaningful and logical. However, there were also instances where GPT-3 appeared to overly generalize its learning about COVID-19 to unseen use cases and often used the

Table 7: Responses by GPT-3 for Task-B

Prompt: Given the context of covid-19 pandemic, explain how a post is offensive or not offensive.
(1) I think Coranavirus waited until we were stupid enough to attack. Like it saw people eating tide pods and was like now's My chance. –[Label: maladaptive]
Response: <i>This post is considered <u>offensive</u> as it implies that the covid-19 pandemic is a consequence of people being stupid, which is generally considered an insensitive thing to say. Additionally, it implies that the virus is intelligent and has malicious intent, which is not the case.</i>
(2) My uncle stopped smoking because of coronavirus. RIP uncle Jim. –[Label: maladaptive]
Response: <i>This post is <u>not offensive</u>. It is a tribute to someone who has passed away due to the consequences of COVID-19. It also serves as a reminder that smoking can be dangerous and can lead to serious health complications.</i>
(3) When the 2020 lockdowns started, people bought all the toilet paper for their a**holes. Well, that's what he claimed. They actually bought it for they're a**holes. – [Label: adaptive]
Response: <i>This post is <u>offensive</u> because it implies that people are lying about why they stocked up on toilet paper and implies that they are doing something inappropriate with the toilet paper. Additionally, it uses profane language.</i>

phrase ‘trivializes the severity of the coronavirus pandemic’ without clearly stating how. Task-A, Input-1 is one such case where even *diapers* were perceived to be a serious measure for protection against COVID-19. The explanation generated for Task A, Input-2 seems to be very apt at the first glance. However, it does not connect the ‘airborne’³ concept which ascribed the humorous angle to the text. The absence of offensive words leads to false prediction to adaptive class as in Task-B, Input-2. Whereas the presence of profane or rude language has led to classification as maladaptive as in Task-B, Input-3.

To evaluate the quality of explanations generated for Task A and Task B, we asked a team of three human annotators to label each explanation on a Likert scale of 1 to 5. The label 1 stands for ‘completely incorrect explanation’ and 5 stands for ‘perfectly correct explanation’. We considered the average ratings by three human annotators for our analysis. For **Task-A**, 68 samples out of 100 samples received a rating of more than 3 indicating a reasonably good explanation. However, only 60% of humorous samples received a score of above 3 for explanations generated by GPT-3 whereas 80% of non-humorous samples received a score of above 3. For **Task-B**, 71 samples out of 100 samples received a rating of more than 3. However, only 42.4% of maladaptive samples received a score of above 3 whereas 85% of adaptive samples had an explanation with a score of more than 3.

³Boeing 737 Max Groundings: https://en.wikipedia.org/wiki/Boeing_737_MAX_groundings

5 DISCUSSION

Topical Humor processing is a formidable task due to the required highly endemic knowledge. During the annotation process, we discovered two key reasons behind disagreement: (a) the annotator is not sufficiently aware of the event and therefore, fails to see the humor, and (b) the annotator finds the material insensitive or even offensive and therefore, does not find it funny. Likewise, for the task of maladaptive humor annotation, it is often the case that only one or two annotators may label a post as maladaptive and the majority of the annotators label it as adaptive. Point (a) is clearly a case of insufficient knowledge and can be ameliorated by recruiting qualified annotators. Point (b) is due to subjectivity and the popular approach is to take *majority vote* for creative tasks where the differences are bound to emerge [23, 29]. Having diverse, representative, and sufficiently qualified annotators, therefore, becomes increasingly important. Nevertheless, we feel that it might a good idea to retain the vote percentage so that downstream applications have the flexibility to determine between majority vote (e.g. *humor generation*) or minority vote (e.g. *culture-specific maladaptive humor detection*).

Over 20% of the samples in our dataset is marked as *maladaptive humor* targeting a racial or political group, nationality⁴ or communities such as *k-pop fans*, and *anti-vaxxers*. Consider the Reddit post, ‘What’s the difference between kung pao and the coronavirus?’. One’s Chinese take out, the other takes out Chinese.’ The onion headline “Trump Blames Nation’s Susceptibility To Coronavirus Outbreak On Weakness Of America’s Race-Muddled Gene Pool” clearly brings out the underlying issue of race discrimination. The post “COVID-19’s primarily killing the unvaccinated. The vast majority of which are idiots.” is clearly a case of microaggression against those who are skeptical of vaccination. The detection of these subtle microaggressions against communities is a formidable obstacle in creating a safe and inclusive digital atmosphere.

In our dataset, the majority of the stereotypes are around one’s nationality whether it’s American, Indian, Korean, or any other. Over 21.7% of identified stereotypes belong to the race category followed by the Chinese Population (20.3%) and nationality (20.3%). For instance, the intelligence of US citizens was made fun of through jokes such as “If there’s one thing Americans can look forward to when the pandemic ends. Is that the average IQ of the population will increase.” Another example is ‘Ohio lawmaker asks if “**race**⁵ population” is hit hard by coronavirus because they don’t wash their hands as well’. Interestingly, old political policies were also used. One such example is “What do they call the Coronavirus in China?. The One Grandparent Policy” that takes a jibe at the One Child policy of China. The stereotypes were not necessarily offensive. However, we can imagine the spread of information required to adequately handle hate speech disguised as humor. In addition, the identified stereotypes could help in the task of norms or stereotype detection [13, 14] and identifying *unfair/harmful* stereotypes.

In addition, humor facilitate quick information sharing [38]. During COVID-19, a significant number of memes and textual humor

⁴“COVID-19 Has Driven Racism And Violence Against Asian Americans: Perspectives From 12 National Polls”, Health Affairs Forefront, April 12, 2022.

⁵Race is not mentioned to avoid further stereotyping

were indirectly regulating or condoning people's behavior. Consider the text, 'In the quarantine, I see you.. Without the quarantine, ICU'. These humor pieces could be used to develop succinct yet delightful information campaigns for public welfare.

5.1 Limitations

Humor is a complex creative construct with high subjectivity. The expressed comic content within a humorous text may evoke feelings of mirth or confusion in some and repulsion in others. Moreover, the use of certain words such as sh*t, fu*k that are used to express surprise, shock, or disbelief may be offensive to some as they are considered to be a part of 'rude language'. For example, Input (2) in Table 7 could be perceived as harmless humor whereas others may find it insensitive and offensive since the text deals with the death of a person. Humor annotation thus becomes a huge challenge due to an annotator's own belief system and the awareness of the world around them. We openly acknowledge that the difference in opinion may occur in the annotated posts.

5.2 Ethical Concern and Broader Impact

Human annotators were college graduates, all of whom identify as female. Keeping this in mind, the annotation may have a slight bias in terms of gender, age, or geographical and cultural representation.

The data obtained from Reddit social media platform via the API are guaranteed to be anonymized. This ensures that the privacy of the authors of the humorous and non-humorous posts is respected and not revealed. As this work studies the different maladaptive humor styles in the posts obtained, there may be offensive or potentially harmful content such as racist, sexist, ageist, trans/homophobic comments, comments insensitive to religious views, political views present in the dataset, which are essential to demarcate these offensive jokes from the adaptive ones. We believe that this study will highlight the offensive speech masked under humor and encourage more research in this direction.

6 CONCLUSION AND FUTURE SCOPE

Creating a reliable yet diverse dataset of humor in any language is a challenge. In this paper, we presented a well-curated dataset of COVID-19 humor tagged with a variety of relevant labels. We believe that this dataset will enable research on the topical humor generated during the pandemic and its role in *information sharing* or even as a *behavior regulator*. The GPT-3 generated promising explanations and is able to identify the pun and word-play style of humor. However, there is significant scope for improvement in maladaptive humor detection and identification of covert toxicity in syntactically non-toxic speech. Online moderation is a highly sensitive zone and moderation shall be performed with caution. The profanities in the text and the use of rude language are one such dilemma. We believe that the theme and the target labels can help us better understand when it is acceptable to be offensive.

REFERENCES

- [1] Vikram Ahuja, Radhika Mamidi, and Navjyoti Singh. 2018. From Humour to Hatred: A Computational Analysis of Off-Colour Humour. In *CCF International Conference on Natural Language Processing and Chinese Computing*. Springer, 144–153.
- [2] Umair Akram, Kamila Irvine, Sarah F Allen, Jodie C Stevenson, Jason G Ellis, and Jennifer Drabble. 2021. Internet memes related to the COVID-19 pandemic as a potential coping mechanism for anxiety. *Scientific reports* 11, 1 (2021), 1–8.
- [3] Patrizia Amici. 2020. Humor in the age of COVID-19 lockdown: An explorative qualitative study. *Psychiatria Danubina* 32, suppl. 1 (2020), 15–20.
- [4] Luca Bischetti, Paolo Canal, and Valentina Bambini. 2020. Funny but aversive: a large-scale survey of the emotional response to COVID-19 humor in the Italian population during the lockdown. (2020).
- [5] Axel Bruns and Jean Burgess. 2015. Twitter hashtags from ad hoc to calculated publics. *Hashtag publics: The power and politics of discursive networks* (2015), 13–28.
- [6] Karyn Buxman. 2012. Types of Humor: The Good, The Bad, and The (Sometimes) Ugly. URL: <https://aath.memberclicks.net/assets/docs/HumorResources/the%20goodkaryn%20aat%20hr.pdf> (2012).
- [7] Simon Critchley. 2011. *On humour*. Routledge.
- [8] Christie Davies. 2008. Undertaking the comparative study of humor. *The primer of humor research, Berlin: Mouton de Gruyter* (2008), 157–182.
- [9] Lambert Deckers and Philip Kizer. 1975. Humor and the incongruity hypothesis. *The Journal of Psychology* 90, 2 (1975), 215–218.
- [10] David JA Dozois, Rod A Martin, and Peter J Bieling. 2009. Early maladaptive schemas and adaptive/maladaptive styles of humor. *Cognitive therapy and research* 33, 6 (2009), 585–596.
- [11] Federica Durante and Susan T Fiske. 2017. How social-class stereotypes maintain inequality. *Current opinion in psychology* 18 (2017), 43–48.
- [12] Marta Dynel. 2021. COVID-19 memes going viral: On the multiple multimodal voices behind face masks. *Discourse & Society* 32, 2 (2021), 175–195.
- [13] Maxwell Forbes, Jena D Hwang, Vered Shwartz, Maarten Sap, and Yejin Choi. 2020. Social chemistry 101: Learning to reason about social and moral norms. *arXiv preprint arXiv:2011.00620* (2020).
- [14] Yi R Fung, Tuhin Chakraborty, Hao Guo, Owen Rambow, Smaranda Muresan, and Heng Ji. 2022. NormSAGE: Multi-Lingual Multi-Cultural Norm Discovery from Conversations On-the-Fly. *arXiv preprint arXiv:2210.08604* (2022).
- [15] Ahmed T Hussein and Lina Nabil Aljamili. 2020. COVID-19 humor in Jordanian social media: A socio-semiotic approach. *Heliyon* 6, 12 (2020), e05696.
- [16] Dirk Kranz, Nicole Maria Thomas, and Jan Hofer. 2021. Changes in age stereotypes in adolescent and older participants of an intergenerational encounter program. *Frontiers in Psychology* 12 (2021), 658797.
- [17] Giseline Kuipers. 2002. Media culture and Internet disaster jokes: Bin Laden and the attack on the World Trade Center. *European Journal of Cultural Studies* 5, 4 (2002), 450–470.
- [18] Tiane L Lee and Susan T Fiske. 2006. Not an outgroup, not yet an ingroup: Immigrants in the stereotype content model. *International Journal of Intercultural Relations* 30, 6 (2006), 751–768.
- [19] Dafna Lemish and Nelly Elias. 2020. "We decided we don't want children. We will let them know tonight": Parental humor on social media in a time of coronavirus pandemic. *International Journal of Communication* 14 (2020), 27.
- [20] Keith B Maddox. 2006. Rethinking racial stereotyping, prejudice, and discrimination. *Psychological Science Agenda* 20, 4 (2006).
- [21] Rod A Martin, Patricia Puhlik-Doris, Gwen Larsen, Jeanette Gray, and Kelly Weir. 2003. Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of research in personality* 37, 1 (2003), 48–75.
- [22] Ariadna Matamoros-Fernández, Aleesha Rodriguez, and Patrik Wikström. 2022. Humor That Harms? Examining Racist Audio-Visual Memetic Media on TikTok During Covid-19. *Media and Communication* 10, 2 (2022), 180–191.
- [23] JA Meaney, Steven Wilson, Luis Chiruzzo, Adam Lopez, and Walid Magdy. 2021. Semeval 2021 task 7: Hahackathon, detecting and rating humor and offense. In *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)*. 105–119.
- [24] Bertolt Meyer and Frank Asbrock. 2018. Disabled or cyborg? How bionics affect stereotypes toward people with physical disabilities. *Frontiers in psychology* (2018), 2251.
- [25] Trenton D Mize and Bianca Manago. 2018. The stereotype content of sexual orientation. *Social Currents* 5, 5 (2018), 458–478.
- [26] John Morreall. 2012. *Philosophy of humor*.
- [27] Andrew R Olah and Thomas E Ford. 2021. Humor styles predict emotional and behavioral responses to COVID-19. *Humor* 34, 2 (2021), 177–199.
- [28] Ana Laura Pérez. 2021. "The 'hate speech' policies of major platforms during the COVID-19 pandemic". *UNESCO, Document Code: MTD/CI/2021/PI/0/REV1* (2021), 32.
- [29] Sunny Rai, Shampa Chakraverty, Devendra K Tayal, Divyanshu Sharma, and Ayush Garg. 2019. Understanding metaphors using emotions. *New Generation Computing* 37 (2019), 5–27.
- [30] Abira Reizer, Yifat Munk, and Lotem Katz Frankfurter. 2022. Laughing all the way to the lockdown: On humor, optimism, and well-being during COVID-19. *Personality and Individual Differences* 184 (2022), 111164.
- [31] Rohit Revi et al. 2014. Understanding obscenity and offensive humour. *The European Journal of Humour Research* 2, 3 (2014), 98–114.

- [32] Bettina Spencer. 2019. Stereotyping and political decision making. In *Oxford Research Encyclopedia of Politics*.
- [33] Andrea Strinić, Magnus Carlsson, and Jens Agerström. 2021. Occupational stereotypes: Professionals warmth and competence perceptions of occupations. *Personnel Review* (2021).
- [34] Joel Mayo Torres, Leila M Collantes, Emily T Astrero, Arceli R Millan, and Carlo M Gabriel. 2020. Pandemic humor: Inventory of the humor scripts produced during the COVID-19 outbreak. *Torres, JM, Collantes, LM, Astrero, ET, Millan, AR, & Gabriel, CM (2020). Pandemic humor: Inventory of the humor scripts produced during the COVID-19 outbreak. The Asian EFL Journal* 7, 3.1 (2020), 138–164.
- [35] Simon Vurayai. 2020. The paradox of disparagement humor: An analysis of WhatsApp jokes on women in the fight against the Covid-19 pandemic. *Gender & Behaviour* 18, 3 (2020), 15949–56.
- [36] Wikipedia contributors. 2022. Topical humor — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Topical_humor&oldid=1096195133. [Online; accessed 26-November-2022].
- [37] Massih Zekavat. 2021. Employing satire and humor in facing a pandemic. *Humor* 34, 2 (2021), 283–304.
- [38] Xiaoquan Zhao, Maria L Roditis, and Tesfa N Alexander. 2019. Fear and humor appeals in “The Real Cost” campaign: Evidence of potential effectiveness in message pretesting. *American journal of preventive medicine* 56, 2 (2019), S31–S39.

A ANNOTATION TASK

A.1 Annotation Guidelines

The annotators were asked to follow the below steps to determine if a text is humorous:

- Read the entire text-discourse to establish a general understanding of the meaning.
- Determine if the text has implicit/explicit meanings of incongruous / conflicting / ambiguous / unexpected or strong emotions that make text humorous in the given context.
- Determine whether the text invokes a humorous reaction from you.

A.2 Hygiene Test

To ensure that the annotators understand the task and definitions, a sample of pre-labelled 30 posts (hygiene sample) was provided for annotation initially. Annotators were given the unlabelled posts for annotation only when they demonstrate clear understanding of the task over the hygiene sample.

A.3 About Annotators

Initially, a team of eight annotators were given the Hygiene test to determine their understanding of the task. Finally, a team of five annotators were picked for the annotation task. All were pursuing undergraduate degree, with ages in the range of 18-21. All 5 of whom identify as female. All were highly proficient in English. No monetary incentive was provided for annotation.

A.4 Definitions

In this section, we provide the definitions for labels assigned during the data annotation.

Table A1: Key Terms in Annotation

Term	Definition
<i>Humor Style</i>	As provided in Table A3
<i>Theme of humor</i>	Key concept or idea on which the humor is based
<i>Target of humor</i>	a community/individual being targeted in the joke (who are being made fun of)
<i>Stereotype</i>	a generalized belief about a particular category of people

Table A2: Humor Type

Type	Definition
Adaptive (1)	When someone uses humor that’s affiliative or self-enhancing in nature, it’s called adaptive humor. [10]
Maladaptive (0)	When someone uses aggressive or self-defeating humor, it’s called maladaptive humor. [10]

Table A3: Humor Styles

Type	Definition
Satire (S)	criticizing people or ideas in a humorous way, especially in order to make a political point, or a piece of writing that uses this style.
Self-defeating (SD)	excessive self-disparaging humor; amusing others at one's own expense to facilitate integration or gain approval; allowing oneself to be the focus of others' humor [21]
Dark humor (D)	a gruesome or a tragic theme such as death/illness is mixed with an innocuous topic which creates shock and inappropriateness [1]
Pun (P)	Puns are usage of certain words in speech to suggest two or more meanings or different associations of a word, or using two or more words with different meanings that sound the same or nearly the same in order to produce humor.
Sarcasm (SR)	convey thinly veiled disapproval humorously. It is something that mocks or ridicules, usually by saying the opposite of what is actually meant.
Irony (IR)	humor that involves incongruity and disagreement between what is said and what is understood or what is said and what actually happened.
Exaggeration (E)	a statement that represents something as better or worse than it really is, i.e. an overstatement of the truth.
Relief (RE)	when the main purpose of the joke is to act as an outlet for built-up tension by finding humor in morally problematic, dehumanizing conjectures [31]
Superiority (SU)	expresses feelings of superiority or dominance over other people/communities or over a former state of ourselves [26]
Matter of-fact (M)	to introduce a statement in an unemotional and practical way, that gives more details about what has just been said, an explanation of it, or something that contrasts with it.

Table A4: Stereotypes

Stereotype	Definition
Gender (G)	A gender stereotype is a generalized view or preconception about attributes or characteristics, or the roles that are or ought to be possessed by, or performed by, women/men.
Race (R)	Racial categorization reflects the process of placing people into distinct groups based on variation in phenotypic physical features of the face and body such as skin color, hair color and texture, eye shape, nose width, and lip fullness [20]
Sexual (S)	Stereotypes of gender, sexuality, and a group's status in society combine to influence societal views of sexual orientation groups [25]
Social-class (SO)	Stereotyping people based on their differential access to material, social, and cultural resources, which shapes their lives in important ways is social class stereotyping [11]
(Dis)ability (D)	People with physical disabilities face generally negative stereotypes toward them; society tends to view physical disability as an inferior state and as a personal misfortune [24]
Age (A)	Age stereotypes are ambivalent. For example, young individuals stereotypically perceive the older generation as warm but incompetent, while older individuals perceive the young generation as open-minded but foolish [16]
Nationality (N)	Perceivers tend to agree on stereotypes of different nationalities, due to their reliance on certain features of the nation, ranging from politics and economics, to religion, to geography, and relational status, cooperating or conflicting, between one's ingroup and the outgroup [18]
Political (P)	Stereotypes to interpret and understand social issues, inform political beliefs, and form impressions of political candidates [32]
Religious (RE)	Stereotypical views characterizing a person solely based on their faith, the religion they practice or the entity they believe in.
Occupational (O)	Stereotypes resulting in preferential/detrimental treatment based on the social status of the occupation, the type of job/job tasks [33]
Chinese population (C)	Stereotypes specific to the Asian community of the Chinese population, especially during the pandemic based on baseless evidence/claims.

Table A5: Sub-Categories for Maladaptive Humor

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009

Category	Definition
Teasing (T)	the act of persistently annoying someone, especially playfully and with silly jokes about them [6]
Insulting (I)	offensive insults directed to a person or a group [1]
Sexist (X)	suggesting that the members of one sex are less able, intelligent, etc. than the members of the other sex, or referring to that sex's bodies, behaviour, or feelings in a negative way.
Racist (R)	having the belief that some races of people are better than others or having general beliefs about other people based only on their race.
Age-ism (A)	comprises of jokes that stereotype, prejudice, and discriminate against people on the basis of their age.