

Identifying COVID-19-related stigmatizing language against people of Asian descent on Twitter

In this dataset, we have provided a set of tweets that contain covid-related discussion. The task of the labellers is to read each tweet carefully and identify whether it has a stigmatizing language stemming from the COVID19 pandemic, against the Asian population. After the assessment, each tweet should get one of the following labels: a) stigmatizing (low) b) stigmatizing (medium) c) stigmatizing (high) d) not stigmatizing e) unknown/irrelevant

Please refer to the instructions and examples below to know about the details of evaluating each tweet.

Description of the Tasks

Task 1

Psychologists define stigma as follows:

“Stigma messages bear the following attributes: they provide cues: (a) to distinguish people, (b) to categorize distinguished people as a separate social entity, (c) to imply a responsibility for receiving placement within this distinguished group and their associated peril, and (d) to link this distinguished group to physical and social peril.”

Based on this definition, we have prepared a set of instructions for you to detect the tweets containing stigmatizing messages against the Asian population.

Identify whether the full text of the tweet is stigmatizing or not following these steps, and identify the level of stigma in each tweet.

- 1) For the tweets with too little information, please mark them as **unknown/irrelevant**. Potential instances could be tweets with a small comment on a linked article, or simply a retweet of an article, or a reply tweet that is hard to understand without the context.

Example: (a) Tweets: "Scary but it has to be said. Everyone needs to listen. <https://t.co/HxuPgkqRKs>". Here the body of the tweet is not related to our subject, and a decision cannot be made without looking at the linked article. In situations like this, mark the tweet as **unknown/irrelevant**.

2) For the tweets that are unrelated to at least one of COVID or the Asian population topics, please mark them as **unknown/irrelevant**

Example: i. Tweet: "What a great smart and warm hat!". Here the tweet is about a topic other than COVID or people of Asian descent. So it should be marked as **unknown/irrelevant**

3) For the tweets that are relevant, please follow the following steps:

a) Identify whether the tweet is targeting people of Asian descent and/or the Chinese population, either explicitly or implicitly.

- If the particular group of people(i.e., Asians/Chinese) is identified, **move to Step 3b to check for the next criterion**
- If no particular group is targeted, mark the tweet as **not stigmatizing**

Examples: i. Tweet: "Chinese people are responsible for COVID"

Here, Chinese people have been targeted. So the tweet **might be stigmatizing. Move to Step 3b to check for other criteria.**

ii. Tweet: "COVID is a curse. We want to get rid of it"

This tweet is **not stigmatizing** because no particular group of people is mentioned.

b) Identify whether the targeted group is separated from others. If not, mark it as **not stigmatizing**.

- If the group is separated from others, **move to 3c to check for the next criterion**
- If the group is not separated, mark the tweet as **not stigmatizing**

Examples: i. Tweet: "Chinese people are responsible for COVID"

Here, only Chinese people are blamed. So the tweet **might be stigmatizing. Move to 3c to check for the next criterion.**

ii. Tweet: "Human Greed is responsible for COVID. There is a rise in such greed everywhere in the world: China, USA, Europe, India, everywhere". No group is separated here. So the tweet is marked as **not stigmatizing**.

c) Identify whether the user is considering the targeted group as people who may pose a physical or symbolic threat to them.

- If they attribute the targeted group as people who might pose a threat, **move to 3d to check for the next criterion**.
- If they do not associate the targeting group with a potential physical or symbolic threat, mark it as **not stigmatizing**.

Examples: i. Tweet: " In my opinion, people of East Asian countries took the quarantine very seriously". Here, although individuals of Asian descent are separated from others, they are not accused of posing a physical threat to society. So the tweet is marked as **not stigmatizing**.

ii. Tweet: "The Asians are coming to the US and taking away our jobs! and they bring us all sort of diseases! They should go back to where they came from." Here the decision is that the targeting group poses threat to society. So the tweet **might be stigmatizing. Move to 3d to check the next criterion**.

d) then, identify whether the user is blaming the group for the threat. The blame could be either explicit or implicit; if not, mark as **not stigmatizing**.

- If the tweet is blaming the targeting group, **move to 3e**.
- If the targeting group is not blamed for the threat, mark it as **not stigmatizing**.

Examples:

i. Tweet: "Seeing people blaming Chinese for eating everything (wild exotic animals) that causes Wuhan Virus.. are you even human. Some people said they need god and religion.. shouldn't you be praying for them bcs you "have" god? Damn humanity.." In this tweet, a group is separated from others, and the

author is associating the group with the threat, but the tweet is not blaming the group for the threat. So this should be marked as **not stigmatizing**

ii. Tweet: "Those disgusting bat-eating people are only bringing us diseases and meddling with our economics. Government should ban them from entering our country". This tweet separates the targeting group, associates them with threats and blames the group. So the tweet **might be stigmatizing. Move to 3e for the final decision.**

e) Finally, If the tweet is targeting a specific group explicitly or implicitly, and is blaming them for a threat/peril (e.g. covid-19, eating wild animals) either with direct or indirect language or if the tweet contains joke/ridicule/sarcasm against a certain population (e.g., Chinese/East Asian), mark as stigmatizing.

4) In the case where the tweet contains potential misinformation about China/Chinese/East Asians:

a) If the Tweet is adapting from news articles (which might be true or false) without any personal opinions towards the Asian population, mark as unknown/irrelevant;

b) If the Tweet is or adapting information/news(which might be true or false) to target China/Chinese/East Asian ethnicity, mark it as stigmatizing.

Examples i. Tweet: "Biological War started By China Virus in September 2019. Economic War started By China Co's in 2000. Military War By China on Land/In Sea will start in August or March. Xi is Dangerous/Mad Dictator. Islam & China are Real Threat to Mother Earth. Time to Unite &Fight." Here, the tweet author is mentioning some information that might be True or False (misinformation). But irrespective of the trueness of the information, the user is targeting China and Islam as certain groups, separating them from others, associating them with a threat, and blaming them for the threat. So the tweet should be marked as **stigmatizing**.

ii. Tweet: "I've heard the origin of the virus is a wet market in Wuhan. Poor hygiene might have started all this." Here, the tweet author is mentioning some information that might be True or False (misinformation). But irrespective of the trueness of the information, the user is not targeting, separating,

associating with a threat, and blaming the Asian/Chinese population. So the tweet should be marked as **Relevant but not stigmatizing**.

In addition to identifying the category of tweets, please rate the degree of stigma embedded in the tweets that are labelled as stigmatizing:

- a) If you labelled the tweet as stigmatizing, choose an option among "low degree", "medium degree", or "high degree" base on your judgement.
- b) Disregard this part if you labelled the tweet as not stigmatizing or unknown/irrelevant.

Example:

i. Low degree stigma: "Chinese people researched and leaked the COVID-19 virus. Intentional or not, they did it. Shameful."

ii. Medium degree stigma:

Tweet: "Ok but actually it's crazy how it spread so fast n to so many people in China all cause the people in Wuhan eat wild animals. Like a whole outbreak cause someone's men eats snakes n bats"

iii. High degree stigma:

Tweet: "Don't you have anything to eat? #ChinesePeople They are fucking cute how can you eat it? Seriously except for those animal they are also eating dog & cat. crazy, if you don't have anything to eat just be vegan #Wuhan bring problem to another"

Task 2

After labelling each tweet as stigmatizing, not stigmatizing, or unknown/irrelevant, please identify the related topics you perceived from the tweet. Check all the related topics.

The topics might be one of the below:

- Wet market & eating habits
- COVID-19 cases

- Bio-weapon
- Wuhan virus lab
- The Chinese government (Chinese communist party - CCP)
- Anti-stigmatizing (talking against hate speech and racism)
- News
- Information or misinformation
- International politics
- Donald Trump
- Chinese Uyghur Muslims
- Other

Task 3 (optional)

If you would like to make a comment about the tweet, or you want to add another related topic that is not listed in task 2, you can mention it here.

Important: If you think the tweet is a duplicate, meaning you have encountered this tweet before, please put the word "duplicate" in this section.

Examples

Tweet	Proposed Label	Related Topics
The truth is I am not racist. I fear what the Chinese culture is doing to the world around me. I am all for trades with Japan, S Korea, any other country. But PLEASE can we CUT TIES WITH CHINA?! #Canada look at what they've	stigmatizing - medium	other

done		
The fact that they haven't halted immigration from China yet because ""muh open borders to stop muh racism"" is fucking horrifying.	stigmatizing - medium	international politics
@realDonaldTrump China worship devil and devil skill is lying. In only Lord Jesus Christ, there is truth.	stigmatizing - high	other
Ur mad because ppl in China want to know where their food came from (wet market), whereas *you* an *intellectual* want the breast of a chicken that lived and died in its own shit, full of antibiotics and hormones and processed by someone making below minimum wage #coronavirus	not stigmatizing	wet market & eating habits anti-stigmatizing
@FinancialTimes China is the biggest threat. The US & Uk ties are not in trouble. The two are not cancelled just because of the 5G decision. Headlines like this are so tiresome.	unknown/irrelevant	international politics
Director of WHO praising China's	not stigmatizing	anti-stigmatizing

response to the virus and saying the actual risks are posed by other countries, but our racist news media will continue hystericizing China & enabling more anti-Asian racism from Westerners in the process		
That's so awesome haha. Fight the boredom China. https://t.co/Ut5vjHNcjP	unknown/irrelevant	other
Trump wants to build a wall to keep out people who hasn't tried to kill us with BIRD FLU OR now CORONAVIRUS. oh and there has been more the Asians have brought in.	stigmatizing - medium	Donald Trump
@susurrou @YesThatAnna so like I said.. all of a sudden it's racist just to be in a pic with an Asian?? #WhoAreTheRacists #ThinkBeforeYouVirtueSignal , go to a hospital,talk to doctors treating corona let's see how you react.. if you live to tell about it, guy handled it stellar and with humour	unknown/irrelevant	other

Just like slippery Sam ring a bell ambo u know the corrupt politician who had China pay his bills and bent over and let them do what they liked	unknown/irrelevant	other
Will you be ok if your government performance like this for #coronavirus outbreak?	unknown/irrelevant	other
The Kenya Aviation Workers Union threatened to lead its members in boycotting work "if KQ does not stop flying to China in the next...	unknown/irrelevant	international politics
The major cities are largely like that, too. Soap is a rare sight and they still have the traditional chinese mindset of traditional medicine and super fresh and obscure foods being healthy, hence the still eating living animals phenomena. China uses tradition for control.	stigmatizing - medium	information or misinformation
China accused of lying over true scale of coronavirus https://t.co/1xxiPEw6X5 via @MailOnline	unknown/irrelevant	other
Not to mention the genocides that were	unknown/irrelevant	international politics

carried out by communist countries, like the USSR and the Holodomir. Or like what Communist China is doing today with the Uyghur Muslims, putting them in LITERAL concentration camps.		
I would like to stand with them in this crisis, but when I remember what they did to Muslims I prefer silence over what happens to them	stigmatizing - low	Chinse Uyghur Muslims
@Teghte @iamtejumorlah They are testing their theory. This is a research work. A biological weapon l China is using to threaten the world and US in particular. HIV/AIDS was created successfully, Ebola virus was tamed by Nigeria, Now Corona virus. Let's watch and see how far it will go	unknown/irrelevant	bio-weapon
Britain,dutch,Canada and Australia too have canceled flights to and from china,and these are countries that have the capacity to at least contain the virus better than Uganda...but okay lets press the default setting and blame the government https://t.co/lv3dSe5Tmm	not stigmatizing	information or misinformation
China probably want to fuck with Japan	stigmatizing - medium	international politics

2020 Olympics. They hate Japan so much they're probably willing to sacrifice a few millions just to spite Japan.		information or misinformation
We do not know everything about that evil experiment. I heard it was stolen from a lab in Canada. A couple got caught but it got to China any way. That was in the news. This was when I first heard about it. But who knows.	not stigmatizing	information or misinformation bio-weapon
More cases of people infected with the coronavirus have been reported in Lagos and Abuja https://t.co/L6KfrnluW1 #Covid19Out #Nigeria #COVID19NIGERIA #COVID19	not stigmatizing	COVID-19 cases
Colombia reports 320 new cases and 10 new deaths bringing total confirmed cases there to 4,881 and 225 total deaths. #CoronaVirus #Covid19 #CoronaVirusOutbreak	not stigmatizing	COVID-19 cases

It will take all of us to defeat the #covid19 pandemic. Let's stand as one and urge world leaders to join a global pandemic response plan. #ONEWorld https://t.co/hGvKGnzfUn	not stigmatizing	other
Chinese are species of Bats, Pangolins, pigs, snake. They eats all wild animals. Is Chinese are human being? Chinese president and Communist ideology #the Corona virus.	stigmatizing - high	wet market & eating habits
"Chinese eat all kind of wild animals is disgusting, including bats. BUT, this #coronavirus is not just simply caught from bats. It's leaked from the P4 Lab in #Wuhan, which means that it's a human disaster. https://t.co/TwcSZEcrAp "	stigmatizing - low	wet market & eating habits information/misinformation
@BelindadeLucy @pennylongstckng @jaglancy The bat soup story was the v1 propaganda cover story from the CCP, before they switched to blaming America. It wasn't bats in a market. It wasn't America. It was a leak from the lab where they were making GM corona viruses for published research.	not stigmatizing	The Chinese Government (CCP) bio-weapon

<p>The amount of racist posts I see about Asians eating cats, dogs, bats and more these days are really upsetting. Racism was never okay. You can't use a pandemic to try and normalise it. Stop this disgusting and hurtful behaviour.</p> <p>#Asian #Racism #Coronavirus #COVID19 #NotOkay</p>	not stigmatizing	anti-stigmatizing wet-market and eating behaviour
--	------------------	--

Labels

Unknown/irrelevant

not stigmatizing

Stigmatizing - low degree

Stigmatizing - medium degree

Stigmatizing - high degree

Description of the Labels & Examples

1 - Unknown/Irrelevant: The tweet is off-topic or does not provide enough information/context for coders to identify whether it's stigmatizing or not

2 -not stigmatizing: The Tweet does not contain stigmatizing language towards Asians

3 - Stigmatizing - low degree: The Tweet contains stigmatizing language towards Asians, and you find the degree of stigma in the tweet low

4 - Stigmatizing - medium degree: The Tweet contains stigmatizing language towards Asians, and you find the degree of stigma in the tweet medium.

5 - Stigmatizing - high degree: The Tweet contains stigmatizing language towards Asians, and you find the degree of stigma in the tweet high

