

Multilingual Emoticon Prediction of Tweets about COVID-19 😊

Stefanos Stoikos

Pomona College

st.stoikos@gmail.com

Mike Izwicki

Claremont McKenna College

mike@izwicki.me

Abstract

Emojis are a widely used tool for encoding emotional content in informal messages such as tweets, and predicting which emoji corresponds to a piece of text can be used as a proxy for measuring the emotional content in the text. This paper presents the first model for predicting emojis in highly multilingual text. Our BERTmoticon model is a fine-tuned version of the multilingual BERT model (Devlin et al., 2018), and it can predict emojis for text written in 102 different languages. We trained our BERTmoticon model on 54.2 million geolocated tweets sent in the first 6 months of 2020, and we apply the model to a case study analyzing the emotional reaction of Twitter users to news about the coronavirus. Example findings include a spike in sadness when the World Health Organization (WHO) declared that coronavirus was a global pandemic, and a spike in anger and disgust when the number of COVID-19 related deaths in the United States surpassed one hundred thousand. We provide an easy-to-use and open source python library for predicting emojis with BERTmoticon so that the model can easily be applied to other data mining tasks.

1 Introduction

The COVID-19 pandemic has caused intense emotional reactions on social media. Some tweets are sad:

This Corona stuff is no joke. Watching people get laid off at work today really made me open my eyes. Wish it was all over. 😭😢🙏

And other tweets are angry:

I saw that bottles of Purell were selling for \$149! Go away price gougers and go away coronavirus! 😡😡😡😡😡😡

What both of these tweets have in common is that their emotional content is captured by emojis present in the tweet's text. Emojis are often used in informal tweets sent between friends (Danesi, 2016), but most tweets do not contain emojis. For example, the following tweet by the BBC (a major British newspaper) is clearly meant to help us find joy amidst the stress of COVID-19:

Father dresses as Transformers character Bumblebee to surprise his son on his first day back at school after lockdown.

But there are no emojis in the text to indicate that the tweet is joyful. One possible emoji for this tweet would be the “grinning face” (😊), but more subtle emojis like “grinning face with tongue” (😋) or “grinning face with smiling eyes” (☺) would also be appropriate and convey slightly different emotions. The goal of this paper is to automatically annotate these emoji-less tweets with appropriate emojis in order to better understand the emotional content in online discussions of COVID-19.

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>.

language	tweet	emoji prediction (top 5)
English	Washington sick man Is 1st in US to Catch Newly Discovered Dangerous Pneumonia. Get out your face masks folks! #coronavirus #wuhan #mask	😂🤣😊⚠️👤
French	Un homme malade de Washington est le premier aux États-Unis à attraper une pneumonie dangereuse nouvellement découverte. Sortez vos masques! #coronavirus #wuhan #mask	😂🤣😎⚠️👤
German	Der kranke Mann aus Washington ist der erste in den USA, der an einer neu entdeckten gefährlichen Lungenentzündung erkrankt. Holen Sie sich Ihre Gesichtsmasken Leute! #coronavirus #wuhan #mask	😂🤣😊😎⚠️
Hebrew	אוללה בושינגטון הוא הראשון בארה"ב כדי לTCPOM דלקת ריאת מסוכנת שהתגלתה לאחרונה. תזעאו את המסוכנות פניהם שלכם, חבר'ה! נגף הקורונה והוא מסכה	😂🤣😎⚠️😡
Indonesian	Orang sakit Washington adalah yang pertama di AS untuk Menangkap Pneumonia Berbahaya yang Baru Ditemukan. Keluarkan masker wajah kalian! #coronavirus #wuhan #mask	⚠️👤😊⚠️⚠️
Italian	Un malato di Washington è il primo negli Stati Uniti a contrarre una polmonite pericolosa scoperta di recente. Tira fuori le tue maschere per il viso, gente! #coronavirus #wuhan #mask	😂🤣😊⚠️😡
Japanese	ワシントンの病人は新しく発見された危険な肺炎を捕まえるために米国で最初です。フェイスマスクの人を出してください！ #コロナウイルス #武漢 #マスク	❓👤😊⚠️⚠️
Portuguese	Homem doente em Washington é o primeiro nos Estados Unidos a pegar pneumonia perigosa recém-descoberta. Tirem suas máscaras, pessoal! #coronavirus #wuhan #mask	😂🤣😊⚠️⚠️
Spanish	Un enfermo de Washington es el primero en los Estados Unidos en contraer una neumonía peligrosa recién descubierta. ¡Saquen sus mascarillas, amigos! #coronavirus #wuhan #mascara	⚠️👤😊⚠️⚠️
Tagalog	Ang taong may sakit sa Washington ay Ika-1 sa Estados Unidos upang Makuha ang Bagong Nakatuklas na Mapanganib na pneumonia. Lumabas ang iyong mga maskara sa mukha mga kamag-anak! #coronavirus #wuhan #mask	⚠️👤😊⚠️⚠️

Figure 1: BERTmoticon predicts good emojis in a wide variety of languages. All non-English text above was translated from the English using Google Translate.

Prior work on predicting emojis from the text of a tweet (Barbieri et al., 2017, Felbo et al., 2017, Zhang et al., 2019) has focused only on English language tweets. Models submitted for the SemEval 2018 Task 2 (Barbieri et al., 2018) are the most multilingual emoji prediction models currently published, but this task considered only English and Spanish tweets. Because COVID-19 is a worldwide phenomenon, however, to understand emotional responses to COVID-19, we must be able to predict emojis in all languages used on Twitter. We therefore introduce the first highly multilingual model for emoji prediction, which we call BERTmoticon. Our model is based on fine-tuning the multilingual BERT model (Devlin et al., 2018), which was trained on a dataset of 102 distinct languages. Figure 1 shows the output of our model on a tweet translated into ten different languages.

A large body of work has emerged analyzing tweets about COVID-19. One prominent line of research attempts to identify how misinformation about the disease spreads online (Elhadad et al., 2020, Kouzy et al., 2020, Prabhakar Kaila et al., 2020, Sharma et al., 2020, Yang et al., 2020). An important subcategory of this research investigates the spread of racist (Budhwani and Sun, 2020, Schild et al., 2020) and ageist (Jimenez-Sotomayor et al., 2020) misinformation. Other research more similar to our own investigates the sentiment of tweets about the coronavirus. Some of this research focuses on specific locations such as Belgium (Kurten and Beullens, 2020), Paris (Saire and Cruz, 2020), Poland (Jarynowski et al., 2020) or India (Das and Dutta, 2020). Other research studies English-language tweets (Rajput et al., 2020, Yin et al., 2020) over wider geographic areas. Our research stands out from this prior work in two important ways. First, we do not consider a subset of tweets about COVID-19, we consider all tweets, written in all languages, sent from anywhere in the world. This is a significantly more challenging technical problem than previous research addressed, but it is also much more useful. Second, we are the first paper to consider the more general emoji prediction problem rather than the sentiment prediction problem. In sentiment prediction, the goal is to assign a positive or negative sentiment to each tweet, and for the coronavirus topic it can be difficult to decidedly assign one sentiment. Tweets about COVID-19 can express negative sentiments because the disease has killed millions of people and forced us to make drastic changes to our lifestyles but also can contain funny, uplifting content exhibiting a positive sentiment such as:

Happy #NationalCatDay from my beautiful adopted pandemic pet!

In the emoji prediction task, we are able to get a more fine-tuned emotional understanding of tweets. For example, we can answer questions like: is the tweet sad? angry? joyful?

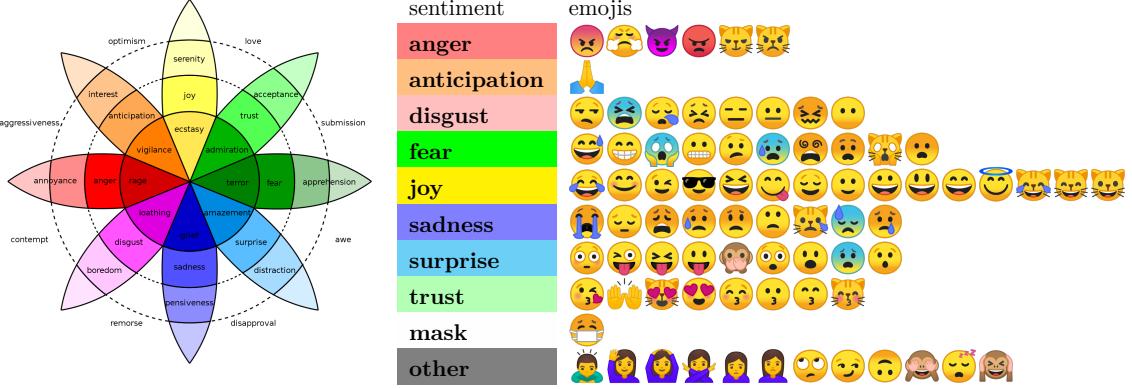


Figure 2: (*Left*) The Plutchik wheel of emotions. (*Right*) We have grouped the emoticons into 10 different categories: 8 emotional categories from the Plutchnik wheel, 1 category for the “face with medical mask” emoticon, and 1 category for all other emoticons that represent emotions that are not clearly in the Plutchik wheel.

Our contributions are as follows. In Section 2 we introduce the first dataset for training highly multilingual emoji prediction models, TwitterEmoticon. We then use this dataset to train the first highly multilingual emoticon prediction model, BERTmoticon. Our model is open source and has an easy to use PyPi package.¹ In Section 3, we introduce the first highly multilingual dataset of tweets about COVID-19, called TwitterCOVID. To generate the dataset, we introduce a novel dataset generation method combining the Twitter API, Bing Translate, and the spaCy tokenization library (Honnibal and Montani, 2017). We then apply the BERTmoticon model to the TwitterCOVID dataset to map how Twitter users across the world have emotionally responded to a variety of COVID-19 news events. This is the first highly multilingual emotion analysis of tweets in any language, and by far the most comprehensive analysis to-date specifically about the COVID-19 pandemic.

2 The BERTmoticon Model

In this section, we first describe the emoticons we are trying to predict and present the TwitterEmoticon dataset that the BERTmoticon model was trained on. Then we describe our training procedure and model evaluation results. We take particular care to ensure that the TwitterEmoticon dataset is sampled from a similar distribution to the TwitterCOVID dataset analyzed in Section 3 below in order to ensure that the BERTmoticon model will transfer well to this unlabeled dataset.

2.1 The Target Emoticons

Emojis were first added to the Unicode standard in 2010, and the current version of the standard (12.1.0) defines 3304 different emojis (The Unicode Consortium, 2019). Prior work on emoji prediction has limited itself to predicting only a subset of the available emojis. For example, Barbieri et al. (2017) consider only the 20 most commonly used emoji, and DeepMoji (Felbo et al., 2017) considers only 64 emoji. There are two primary reasons for only considering a subset of emoji. First, emoji-usage follows a power law distribution where the top 1% of most used emoji account for over 99% of all emoji usage.² There is therefore very little training data for the less popular emoji, and so we cannot expect a classifier to have high prediction accuracy for these emoji. Second, many emoji (e.g. the Greek Flag emoji ☺) do not contain emotional information, and so the ability to predict these emoji does not help us understand the emotional content of text.

We follow previous work and focus on predicting only a limited set of emoji. Specifically, we focus on the original 80 emoji defined in the Unicode standard’s emoticon code block (code points 0x1f600

¹ <https://github.com/Stefanos-stk/Bertmoticon>

² See <http://www.emojitracker.com/> for real-time stats on Twitter emoji usage.

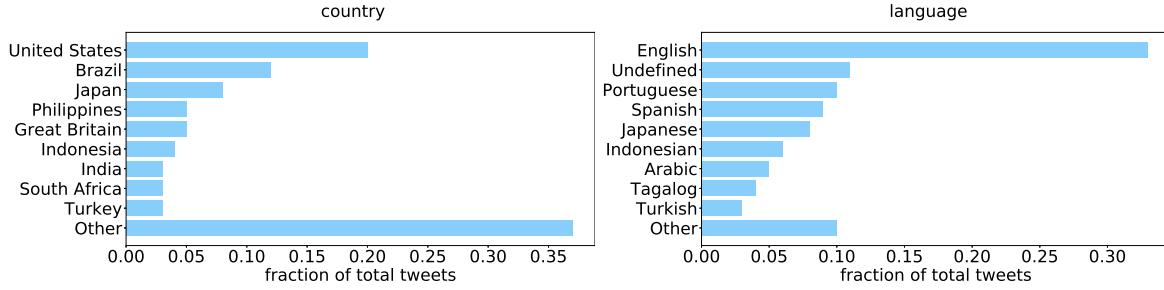


Figure 3: Stats on the most common countries of origin (*left*) and languages (*right*) for tweets in the TwitterEmoticon dataset. Languages are determined using Twitter’s API, which has official support for 66 languages. It is known, however, that more than 100 language are actively used on Twitter (Hong et al., 2011), and our BERTmoticon model supports all of these languages. All prior work on emoji prediction has focused on only 1 or 2 languages.

- 0x1f650). In common usage, the words *emoji* and *emoticon* are interchangeable, but in this paper we adopt the Unicode Standard’s definitions of these terms. By these definitions, an *emoji* is any one of 3304 pictographs that are not part of any written language, and an *emoticon* is one of the original 80 emoji defined in the code block specified above. We limit our analysis to emoticons for three reasons. First, they are the most commonly used emoji on twitter, so we can expect to achieve relatively high accuracy. Second, each emoticon represents an emotion (emoticon is a portmanteau of emotion and icon). Third, the emoticon block contains the “face with medical mask” emoji (😊), which is important for our case study analyzing emotional responses to the coronavirus.

Figure 2 shows the 80 emoticons and a mapping from these emoticons to the Plutchik wheel of emotions (Plutchik, 1991). The Plutchik wheel is a standard psychological model for encoding emotions that has been highly influential in emotion prediction systems (e.g. Kant et al., 2018, Liu et al., 2019, Suttles and Ide, 2013). It has 8 primary emotional categories (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust). These emotions are arranged spatially so that similar emotions (e.g. joy, trust) appear near each other, and dissimilar emotions (e.g. joy, sadness) appear opposite each other. Furthermore, each emotional category is broken down into sub-categories that encode the strength of the emotion (e.g. ecstasy is an extreme form of joy, and serenity is a mild form of joy).

There is currently no standard mapping from emoticons onto the Plutchik wheel, and in Figure 2 (*right*) we provide a suggested mapping. To generate this mapping, we manually assigned each emoticon to an emotion based on the description of the emoticon on the website emojipedia.org. The mapping is not perfect. The category joy has many emoticons representing different facets of joy, but the category anticipation has only a single emoticon. We emphasize that our BERTmoticon model will predict raw emoticons directly, but we present the mapping onto the Plutchik wheel emotions to help make the wide array of emoticon emotions more easily understandable.

2.2 The TwitterEmoticon Dataset

The TwitterEmoticon dataset is designed for training a classifier that takes as input a tweet and outputs an emoticon that represents the emotion of the tweet. To generate the dataset, we collected all geolocated³ tweets sent over the six month period between January and June, 2020. Approximately 400 million tweets meet this criteria. Then we filtered these tweets so that only tweets containing one of our 80 target emoticons were included, and any retweets were removed. In total, the TwitterEmoticon dataset contains 64.2 million tweets sent by 4.2 million users. The tweets are written in 66 different languages and were sent from 246 different countries. Figure 3 shows the total number of tweets per

³ Twitter users can adjust their privacy settings to include different amounts of geolocation metadata. In particular, they can include the exact GPS coordinate that a tweet was sent from, an approximate location (for example, the city that the tweet was sent from), or no location information at all. We say that a tweet is *geolocated* if any of this metadata is included about the tweet. Approximately 1% of all tweets are geolocated.

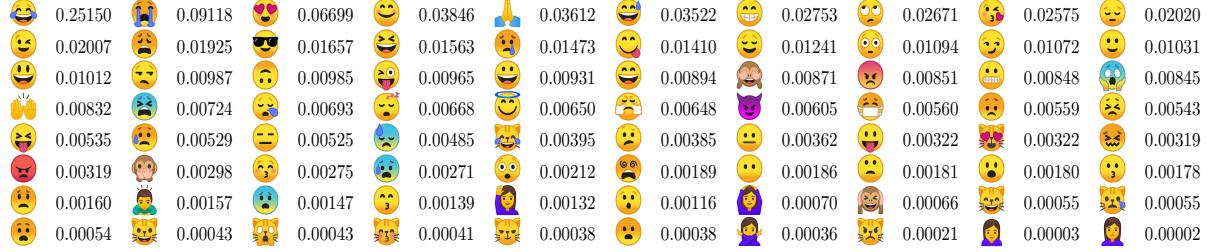


Figure 4: The 80 target emoticons we are trying to predict, and their fraction of all emoticons in the TwitterEmoticon dataset. The distribution follows a power law.

language, and Figure 4 shows the frequency of each emoticon in the dataset.

We preprocess each tweet by replacing all user mentions with a special token `<mention>` and all URLs with a special token `<url>` and deleting all emojis. We decided to keep all hashtags because hashtags can contain potentially valuable emotional content useful for emoticon prediction. Finally, we delete all emoticons from the tweet, and use the emoticons as the tweet’s classification label. Most tweets have only a single emoticon label, but some tweets have multiple emoticons. This is a problem because standard multi-class classification techniques require only a single label per data point. We address the issue by following the procedure established by (Felbo et al., 2017). If a tweet has multiple emoticons, then we duplicate it in the training data once for each emoticon, with each instance being labeled by a single one of the emoticons.

We carefully split the TwitterEmoticon dataset into training, validation, and test sets ensuring that no user is present in more than one set in order to prevent data leakage. In particular we assign 80% of users to the training set, 10% to the validation set, and 10% to the test set. The tweets contained in each set are then the tweets sent by each of the users in the set.

2.3 Training Protocol

Our BERTmoticon model is the multilingual BERT model (Devlin et al., 2018) fine-tuned on the TwitterEmoticon dataset. The multilingual BERT model is a popular model for fine-tuning because it achieves state-of-the-art performance on a wide variety of natural language tasks. It was trained on data from 102 distinct languages, and the language of each training sample need not be known for either training or inference. Followup research has shown that the multilingual BERT model has language-independent internal representations that allow it to encode information from languages it has not seen during training time (Pires et al., 2019, Wu et al., 2019). Feng et al. (2020) recently released a more advanced version of the multilingual BERT model that achieves better performance on standard NLP tasks and uses 109 training languages. We would expect better performance on our emoticon-prediction task using this more advanced multilingual BERT model, but we did not use this model because all of our experiments were completed before this model was publicly released.

We followed a two step fine-tuning procedure. First, we trained only the last layer of the model, generating a model we call BERTmoticon-LL. Then, we trained all parameters to generate the BERTmoticon model, warm starting from BERTmoticon-LL. We used the validation set to select optimal hyperparameters for both models. BERTmoticon-LL was trained using Adam (Kingma and Ba, 2014) with a learning rate of 10^{-4} , and BERTmoticon was trained using Adam with a learning rate of 10^{-5} . Both models used a batch size of 64. A single epoch on the TwitterEmoticon dataset took approximately 6 days to run on one NVidia GeForce RTX 2080 GPU. We trained both models on a single epoch, but found that the model converged before the epoch was finished.

2.4 Model Evaluation

The BERTmoticon-LL model achieves a Macro-F1 score of 0.159 on the test set and the BERTmoticon model achieves a Macro-F1 score of 0.210. Table 1 shows a performance breakdown by emoji and by language. Prediction performance on each language varies dramatically because each

	😂	😍	😊	😊	😊	😊	😊	😊	😭	😊	😊	😊	🙏	Macro-F1
Model														
BERTmoticon-LL	0.440	0.190	0.111	0.185	0.008	0.111	0.011	0.111	0.018	0.049	0.007	0.226	0.159	
BERTmoticon	0.479	0.309	0.211	0.244	0.078	0.249	0.073	0.256	0.215	0.106	0.035	0.361	0.210	
Language														
Arabic	0.637	0.224	0.000	0.000	0.006	0.002	0.000	0.117	0.064	0.000	0.000	0.268	0.319	
Dutch	0.371	0.141	0.027	0.225	0.004	0.364	0.006	0.149	0.068	0.055	0.000	0.201	0.168	
English	0.506	0.295	0.201	0.230	0.034	0.241	0.031	0.235	0.160	0.128	0.006	0.367	0.221	
French	0.418	0.227	0.056	0.109	0.010	0.414	0.008	0.356	0.083	0.180	0.004	0.196	0.208	
Hindi	0.443	0.103	0.016	0.007	0.002	0.027	0.010	0.067	0.025	0.010	0.000	0.717	0.310	
Indonesian	0.442	0.188	0.053	0.065	0.002	0.218	0.027	0.393	0.073	0.036	0.001	0.339	0.203	
Italian	0.430	0.273	0.220	0.205	0.031	0.439	0.071	0.144	0.168	0.074	0.007	0.119	0.181	
Japanese	0.115	0.199	0.323	0.336	0.000	0.003	0.055	0.223	0.268	0.137	0.004	0.142	0.133	
Portuguese	0.380	0.426	0.310	0.004	0.155	0.128	0.183	0.249	0.294	0.092	0.110	0.340	0.177	
Russian	0.375	0.267	0.000	0.028	0.000	0.249	0.000	0.248	0.000	0.000	0.000	0.044	0.122	
Spanish	0.388	0.381	0.183	0.090	0.087	0.423	0.073	0.254	0.185	0.142	0.065	0.305	0.170	
Tagalog	0.509	0.247	0.070	0.051	0.035	0.174	0.002	0.293	0.154	0.035	0.019	0.225	0.208	
Thai	0.271	0.124	0.023	0.042	0.000	0.012	0.000	0.208	0.009	0.012	0.000	0.022	0.078	
Turkish	0.469	0.132	0.055	0.121	0.059	0.085	0.003	0.118	0.023	0.036	0.002	0.392	0.174	
(*) Undefined	0.485	0.127	0.030	0.010	0.008	0.012	0.003	0.042	0.083	0.041	0.000	0.103	0.171	

Table 1: (*top*) F1 scores for the BERTmoticon-LL and BERTmoticon models on 12 selected emojis, and the Macro-F1 incorporating all 80 target emoticons. The full BERTmoticon model offers significantly better performance across all emoji categories. (*bottom*) Performance of the BERTmoticon model broken down by language on 15 selected languages. The Undefined language corresponds to tweets for which the Twitter API was not able to assign a language. Even for these tweets, which are either written in an unsupported language or do not contain a significant amount of text within them, the BERTmoticon model is able to get performance comparable to many officially supported languages.

bacteria, cdc, china, corona, coronavirus, cough, covid, covid-19, covid19, disease, doctor, epidemic, fever, flatten the curve, flu, lockdown, n95, ncov, nurse, outbreak, pandemic, sars-cov-2, sick, sinophobia, social distancing, trump, vaccine, virus, wuhan

Figure 5: The 29 English-language search terms we used to select tweets. These terms were translated using Bing Translate into 72 other languages as part of our multilingual tweet filtering process.

language uses emojis with different frequencies. Arabic language tweets, for example, use the “crying tears of joy” emoticon (😂) frequently, but rarely use the “smiling face with heart eyes” emoticon (😍). The model has therefore learned to favor predictions of this emoticon whenever these predictions are present in the tweet. As Figure 1 demonstrates, this causes the same text translated into different languages to receive different emoji labels. We believe that this is a strength of our model, as different cultures use emojis differently, and our model is able to capture this fact.

3 Coronavirus Case Study

We now apply the BERTmoticon model to understand the emotional response of Twitter users to news about the coronavirus. We first introduce our TwitterCOVID dataset, then we present an analysis of this dataset.

3.1 The TwitterCOVID Dataset

The goal of the TwitterCOVID dataset is to include any geolocated tweet that references COVID-19 in any language. There is currently no standard procedure for generating multilingual datasets of tweets about a topic, so we used the following four step procedure: (1) We generated a list of 29 English-language search terms related to the coronavirus, as shown in Figure 5. The choice of terms was inspired by the terms used in a dataset generated by Chen et al. (2020), but we also removed Twitterisms like “kungflu” which would not translate well into non-English languages. (We discuss in detail the full differences between our dataset and the dataset of Chen et al. (2020) below.) Our search terms include generic words like “china” and “trump” that are not necessarily about the coronavirus, but given the time period we searched over, many tweets including these terms will be about the coronavirus. (2) We then used Bing’s translation API to translate each of these terms into the 72 languages supported by Bing

Afrikaans, Arabic, Armenian, Bulgarian, Catalan, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Gujarati, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Irish, Italian, Japanese, Kannada, Korean, Latvian, Lithuanian, Malayalam, Marathi, Norwegian Bokmål, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Sinhala, Slovak, Slovenian, Spanish, Swedish, Tagalog, Tamil, Tatar, Telugu, Thai, Turkish, Ukrainian, Urdu, Vietnamese

Figure 6: The full list of 54 languages supported by all 3 tools in our processing pipeline (Bing translate, Spacy, and the Twitter API). Tweets in other languages are also included in our TwitterCOVID dataset, but the filtering step is less accurate for these unsupported languages.

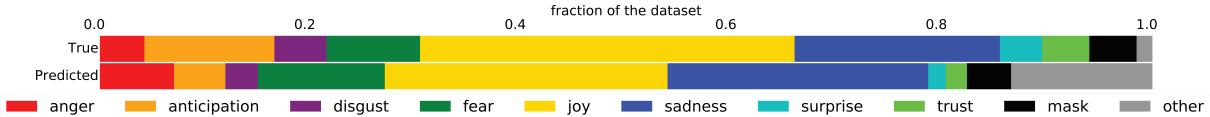


Figure 7: (top) The distribution of emotions in the subset of TwitterCOVID that contain emojis. (bottom) The distribution of emotions in the subset of TwitterCOVID that did not contain emojis, and that we used BERTmoticon to assign predictions for.

translate. (3) We used Python’s spaCy library (Honnibal and Montani, 2017) to tokenize and lemmatize each of the 400 million geolocated tweets sent between January and June 2020. This is the same time period that we examined for the TwitterEmoticon dataset, and so we hope that the BERTmoticon model trained on the TwitterEmoticon dataset will transfer well to this TwitterCOVID dataset. spaCy supports tokenization in 58 different languages, and for each tweet we used the appropriate spaCy module for the language specified in the tweet’s metadata. For languages not directly supported by spaCy, we tokenized on whitespace. (4) Finally, the TwitterCOVID dataset is constructed as the set of all tweets whose lemmatized text contains any of the search terms from the tweet’s language or English. We include both languages in this filtering step because it is common for non-English tweets to use English words like “coronavirus” when referencing the virus. Figure 6 shows the full list of 54 languages that are supported by all 3 services. The TwitterCOVID dataset contains tweets in an unknown number of other languages, and the filtering procedure for these unsupported languages used language-agnostic steps, which likely results in less recall. In total, 16.2 million tweets meet the criteria to be included in the TwitterCOVID dataset.

The other significant dataset of coronavirus related tweets was introduced by Chen et al. (2020). There are two main differences between our dataset and theirs. First, we only include geolocated tweets, whereas they include non-geolocated tweets as well. This results in their dataset being about fifteen times larger than ours, with about 250 million tweets over the same time period. Because their data is not geolocated, however, it is not suitable for understanding how different countries have reacted emotionally to COVID-19. The second difference is that our dataset uses a more advanced language-aware filtering method. They only search for tweets that contain English keywords. Most languages, however, have few words in common with English, and non-Latin based languages frequently do not even use the word “coronavirus” to describe the virus. Chinese tweets, for example, commonly refer to COVID-19 with the string 病毒, and Chinese-language tweets containing this string will get included in our dataset but not in their dataset. As a result of this more advanced processing, the fraction of non-English tweets is much larger in our dataset than theirs (48% versus 38%). Capturing as many non-English tweets as possible about COVID-19 is important for ensuring that our analysis is not unfairly skewed towards English-speaking countries.

3.2 Results

Only 15.11% percent of tweets in the TwitterCOVID dataset contain an emoticon. We used the BERTmoticon model to label the remaining tweets. Figure 7 shows the distribution of tweets present in the dataset vs those we predicted. The anticipation, disgust, joy, surprise, and trust emoticons appear less frequently in the predicted dataset, and the anger, sadness, and fear emotions appear more often in the predicted set. We hypothesize that this difference is due to the fact that more-formal Twitter accounts (such as for newspapers or government organizations) are less likely to use emoji in their tweets, and

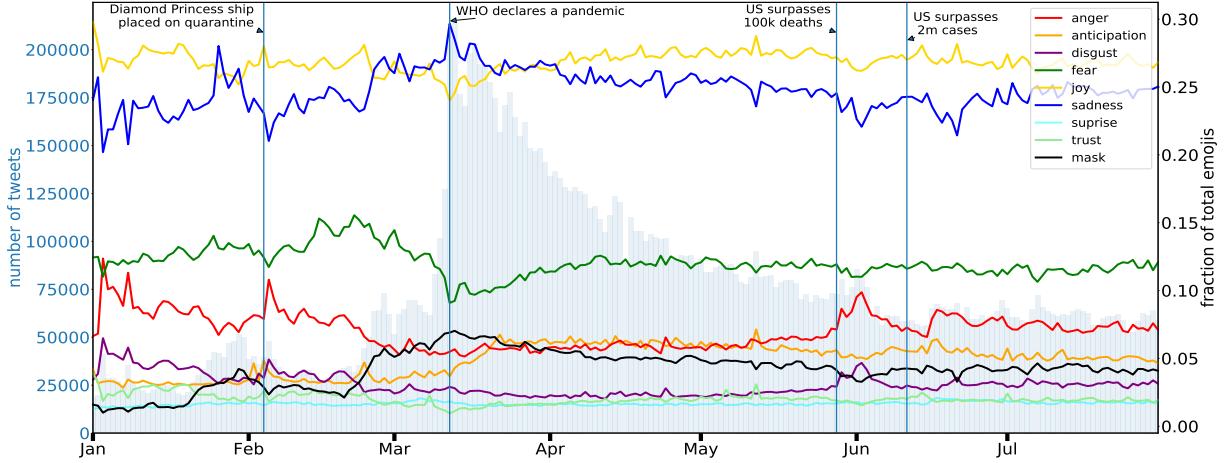


Figure 8: The emotional content of tweets in the TwitterCOVID dataset changes over time and reacts to major news events. The shaded bar plot in the background shows the total number of tweets in the TwitterCOVID dataset sent on a particular day (left y-axis scale), and the colored line charts show the fraction of tweets in a particular day that correspond to each emotion on the Plutchik wheel or the mask emoji (right y-axis scale). We can see clear emotional reactions to the news events labelled with vertical lines.

these formal accounts also tweet about different topics than more informal accounts of ordinary people.

Our main result is shown in Figure 8. For each day, we calculate the fraction of tweets that represent each emotion from the Plutchik wheel (see Figure 2), and we can observe a strong correlation between the emotional content of tweets and important COVID-19 news. For example, on March 11, the World Health Organization (WHO) declared COVID-19 a worldwide pandemic. At the same time, we can see a large spike in tweets about the coronavirus, and in particular we see an increase in sadness and a decrease in joy. Since sadness and joy are at opposite ends of the Plutchik wheel of emotions, it makes sense that a rise in one would cause a fall in the other. As another example, on May 28, the United States had its one hundred thousandth death to the coronavirus. At the same time, we see spikes in anger and disgust. The following tweet from this time period is a representative example:

*How do we tolerate 3000 Americans dying everyday from #COVID19? THREE. THOUSAND.
EVERY. DAY.*

This tweet was not originally sent with any emoticons, but our BERTmoticon model was able to label it with 😔, 😢, 😢, 😷. As another example, notice that the mask emoji usage increases in mid-January. Interestingly, at that time little information was available about COVID-19 and protecting yourself against it. Finally, in early February some important news-events that circulated in Twitter were the Diamond Princess Ship being placed under quarantine and the death of Doctor Li Wenliang (a Chinese doctor who issued a warning about the coronavirus before the pandemic was officially recognized). At that point we notice a spike in anger and disgust.

4 Conclusion

We introduced the BERTmoticon model for multilingual emoji prediction, and used this model to better understand how Twitter users responded emotionally to news about the coronavirus. In follow up studies, we hope to analyze how different countries and language communities reacted differently to events, and have designed our TwitterCOVID dataset and BERTmoticon model to facilitate these cross-sectional analyses. We also hope that the BERTmoticon model will prove useful for analyzing the emotions of text in other contexts outside of COVID-19, and we make the model available in an easy to use Python package to facilitate this process.

References

- Francesco Barbieri, Miguel Ballesteros, and Horacio Saggion. Are emojis predictable? *arXiv preprint arXiv:1702.07285*, 2017.
- Francesco Barbieri, Jose Camacho-Collados, Francesco Ronzano, Luis Espinosa Anke, Miguel Ballesteros, Valerio Basile, Viviana Patti, and Horacio Saggion. Semeval 2018 task 2: Multilingual emoji prediction. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 24–33, 2018.
- Henna Budhwani and Ruoyan Sun. Creating covid-19 stigma by referencing the novel coronavirus as the “chinese virus” on twitter: Quantitative analysis of social media data. *Journal of Medical Internet Research*, 22(5):e19301, 2020.
- Emily Chen, Kristina Lerman, and Emilio Ferrara. Tracking social media discourse about the covid-19 pandemic: Development of a public coronavirus twitter data set. *JMIR Public Health and Surveillance*, 6(2):e19273, 2020.
- Marcel Danesi. *The semiotics of emoji: The rise of visual language in the age of the internet*. Bloomsbury Publishing, 2016.
- Subashish Das and Anandi Dutta. Characterizing public emotions and sentiments in covid-19 environment: A case study of india. *Journal of Human Behavior in the Social Environment*, 0(0):1–14, 2020. doi: 10.1080/10911359.2020.1781015. URL <https://doi.org/10.1080/10911359.2020.1781015>.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2018.
- Mohamed K Elhadad, Kin Fun Li, and Fayez Gebali. Covid-19-fakes: a twitter (arabic/english) dataset for detecting misleading information on covid-19. In *International Conference on Intelligent Networking and Collaborative Systems*, pages 256–268. Springer, 2020.
- Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. *arXiv preprint arXiv:1708.00524*, 2017.
- Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. Language-agnostic bert sentence embedding. *arXiv preprint arXiv:2007.01852*, 2020.
- Lichan Hong, Gregorio Convertino, and Ed H Chi. Language matters in twitter: A large scale study. In *ICWSM*, 2011.
- Matthew Honnibal and Ines Montani. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear, 2017.
- Andrzej Jarynowski, Monika Wojta-Kempa, and Vitaly Belik. Perception of emergent epidemic of covid-2019/sars cov-2 on the polish internet. Available at SSRN 3572662, 2020.
- Maria Renee Jimenez-Sotomayor, Carolina Gomez-Moreno, and Enrique Soto-Perez-de Celis. Coronaviru, ageism, and twitter: An evaluation of tweets about older adults and covid-19. *Journal of the American Geriatrics Society*, 2020.
- Neel Kant, Raul Puri, Nikolai Yakovenko, and Bryan Catanzaro. Practical text classification with large pre-trained language models. *arXiv preprint arXiv:1812.01207*, 2018.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Ramez Kouzy, Joseph Abi Jaoude, Afif Kraitem, Molly B El Alam, Basil Karam, Elio Adib, Jabra Zarka, Cindy Traboulsi, Elie W Akl, and Khalil Baddour. Coronavirus goes viral: quantifying the covid-19 misinformation epidemic on twitter. *Cureus*, 12(3), 2020.

Sebastian Kurten and Kathleen Beullens. # coronavirus: Monitoring the belgian twitter discourse on the severe acute respiratory syndrome coronavirus 2 pandemic. *Cyberpsychology, Behavior, and Social Networking*, 2020.

Chen Liu, Muhammad Osama, and Anderson De Andrade. Dens: a dataset for multi-class emotion analysis. *arXiv preprint arXiv:1910.11769*, 2019.

Telmo Pires, Eva Schlinger, and Dan Garrette. How multilingual is multilingual bert? *arXiv preprint arXiv:1906.01502*, 2019.

Robert Plutchik. *The emotions*. University Press of America, 1991.

Dr Prabhakar Kaila, Dr AV Prasad, et al. Informational flow on twitter–corona virus outbreak–topic modelling approach. *International Journal of Advanced Research in Engineering and Technology (IJARET)*, 11(3), 2020.

Nikhil Kumar Rajput, Bhavya Ahuja Grover, and Vipin Kumar Rathi. Word frequency and sentiment analysis of twitter messages during coronavirus pandemic. *arXiv preprint arXiv:2004.03925*, 2020.

Josimar E Chire Saire and Jimmy Frank Oblitas Cruz. Study of coronavirus impact on parisian population from april to june using twitter and text mining approach. *medRxiv*, 2020.

Leonard Schild, Chen Ling, Jeremy Blackburn, Gianluca Stringhini, Yang Zhang, and Savvas Zannettou. ” go eat a bat, chang! ”: An early look on the emergence of sinophobic behavior on web communities in the face of covid-19. *arXiv preprint arXiv:2004.04046*, 2020.

Karishma Sharma, Sungyong Seo, Chuizheng Meng, Sirisha Rambhatla, Aastha Dua, and Yan Liu. Coronavirus on social media: Analyzing misinformation in twitter conversations. *arXiv preprint arXiv:2003.12309*, 2020.

Jared Suttles and Nancy Ide. Distant supervision for emotion classification with discrete binary values. In *International Conference on Intelligent Text Processing and Computational Linguistics*, pages 121–136. Springer, 2013.

The Unicode Consortium. The Unicode Standard. Technical Report Version 12.1.0, Unicode Consortium, 2019. URL <http://www.unicode.org/versions/Unicode12.1.0/>.

Shijie Wu, Alexis Conneau, Haoran Li, Luke Zettlemoyer, and Veselin Stoyanov. Emerging cross-lingual structure in pretrained language models. *arXiv preprint arXiv:1911.01464*, 2019.

Kai-Cheng Yang, Christopher Torres-Lugo, and Filippo Menczer. Prevalence of low-credibility information on twitter during the covid-19 outbreak. *arXiv preprint arXiv:2004.14484*, 2020.

Hui Yin, Shuiqiao Yang, and Jianxin Li. Detecting topic and sentiment dynamics due to covid-19 pandemic using social media. *arXiv preprint arXiv:2007.02304*, 2020.

Linrui Zhang, Yisheng Zhou, Yang Yu, and Dan Moldovan. Towards understanding creative language in tweets. *Journal of Software Engineering and Applications*, 12:447–459, 01 2019. doi: 10.4236/jsea.2019.1211028.