EmoTag1200 👍: Understanding the Association between Emojis 😄 and Emotions 😍

Abu Awal Md Shoeb
Dept. of Computer Science
Rutgers University
New Brunswick, NJ, USA
abu.shoeb@rutgers.edu

Gerard de Melo
Hasso Plattner Institute
University of Potsdam
Potsdam, Germany
gdm@demelo.org

Abstract

Given the growing ubiquity of emojis in language, there is a need for methods and resources that shed light on their meaning and communicative role. One conspicuous aspect of emojis is their use to convey affect in ways that may otherwise be non-trivial to achieve. In this paper, we seek to explore the connection between emojis and emotions by means of a new dataset consisting of human-solicited association ratings. We additionally conduct experiments to assess to what extent such associations can be inferred from existing data in an unsupervised manner. Our experiments show that this succeeds when high-quality word-level information is available.

1 Introduction

People increasingly rely on digital channels such as mobile instant messaging apps to communicate with their friends, families, colleagues, and communities. Along with this rapid shift in medium, there have been concomitant changes in the way people express themselves in written language (McCulloch, 2019). One notable development has been the emergence of emojis as a new modality, presenting rich possibilities for representation and interaction. Emojis have become ubiquitous in social media and in instant messaging, owing in part to their visual appeal and their ease of use compared to typing out full words on mobile devices.

However, the rise of emojis also substantially appears to stem from their ability to convey affect (Vidal et al., 2016; Zhou et al., 2017). This is evinced by the fact that the most frequently used emojis are smileys and other facial expression symbols that exhibit a direct connection to emotional expression (Ekman and Friesen, 1986). These largely displaced traditional *emoticons* such as ":-)" and ":)", which as well were chiefly used to convey humor and emotion (Derks et al., 2008), as also reflected in their name, a portmanteau of the words *emotion* and *icon*.

This mandates additional analysis of the nexus between emojis and emotion. Past work has compiled a list of sentiment polarity scores for a set of emojis (Novak et al., 2015). Rakhmetullina et al. (2018) categorized a set of 15 emojis into 4 different emotion classes, while Li et al. (2019) used a lexicon-based heuristic to compare connections between emojis and emotions in social media data. Several studies have explored the linguistic connection between words and emojis (Cappallo et al., 2019; Barbieri et al., 2017; Na'aman et al., 2017; Shoeb et al., 2019). However, previous work has not assessed to what extent humans associate particular emotions with different emojis.

In this work, we present EmoTag1200, a dataset of human ratings of association for a set of 150 popular emojis with regard to 8 different emotions. Each of the resulting 1,200 pairs of emojis and emotions has been annotated by 9 human raters on a 5-point scale. The purpose of this endeavor is to measure the degree of emotion that people associate with the use of a given emoji in written expression. As the set of emotions, we consider the eight basic ones in the Wheel of Emotions by Plutchik (1980), i.e., *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, and *trust*. The EmoTag1200 dataset as well as additional emoji-related resources are available online[1].

We assess the emotion scores of this set of emojis and subsequently study a series of simple unsupervised models to predict such emotion intensity scores automatically. For this, we investigate standard pre-trained vector embedding models, but also consider an emoji-centric corpus consisting of 20.8M tweets to study how it can expose semantic relationships between emotion words and

[1] http://emoji.nlproc.org

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 8957–8967,
November 16–20, 2020.

emojis, drawing on additional lexical resources. The results suggest that models drawing on word-level emotion intensity information as background knowledge fare better than vanilla vector embedding models.

2 Background and Related Work

Emotion and Communication. Darwin (1872) was among the first to consider the connection between emotions and their expression in substantial detail. He remarked for instance, that for both animals and humans, anger coincides with eye muscle contractions and teeth exposure, and commented on the fact that humans lift their eyebrows in moments of surprise. His work then goes on to study the role of such forms of facial expression in conveying to others how an animal feels, studying primates as well as human infants and adults.

In light of this important role, humans continue to rely extensively on such nonverbal cues in oral forms of linguistic communication. Although a person's emotion and mood can to some extent be conveyed by means of suitable content words (e.g., "I am happy to hear that!") or interjections ("Wow!"), face-to-face communication has important properties that written communication tends to lack (Bordia, 1997). These include facial expressions of the aforementioned sort, but also gesture and intonation. In certain problem-solving settings, for instance, face-to-face communication may hence prove more efficient and effective (Bordia, 1997).

Accordingly, throughout the history of writing, humans have resorted to surrogate mechanisms to convey emotive signals, attempting to push the boundaries and overcome some of the inherent restrictions of plain written language as a medium, e.g., by means of illustrative embellishments and ornaments (Voronova and Sterligov, 1997). User studies have shown that images (Lang et al., 1999), color (Bartram et al., 2017; Kulahcioglu and de Melo, 2019), and typography (Kulahcioglu and de Melo, 2018, 2020) contribute to conveying affect.

Emoticons. Emoticons such as ":-)" and Japanese 顔文字 (*kaomoji*) such as "(^_^)", both composed from regular symbols, have been in use for several decades. Early studies focused on the use of emoticons in social media. Go et al. (2009) proposed a form of distant supervision by using emoticons as noisy labels for Twitter sentiment classification. Davidov et al. (2010) adopted a similar approach by handpicking smileys and hashtags as tweet labels to train a supervised model to classify the sentiment of tweets.

Emojis. Emoji characters are pictorial, similar to earlier dingbat characters, but also colorful. Despite the lexicographic similarity between the two words *emoji* and *emotion*, etymologically, the former stems from the Japanese words 絵 (*e*, picture) and 文字 (*moji*, character). Emojis originated in Japan in the 1990s and have only recently spread globally. Historically, the spread of emojis has been driven in large part by their adoption in popular messaging and social media platforms, which led, among other things, to their inclusion in Shift JIS, and, subsequently, the Unicode standard. Nowadays, they are ubiquitous in social media and chat applications, but increasingly also in emails and other digital correspondence.

Emojis have a number of different roles. Kaye et al. (2017) explained how emojis may aid the interlocutor in disambiguating utterances that would otherwise remain ambiguous.

One of their principal uses has been to convey emotion, particularly via facial expression emojis, as explained in Section 1. In 2015, Oxford Dictionaries declared the *Face with Tears of Joy* emoji its Word of the Year 2015. Emojis may also be useful as a more instantaneously and widely recognized form of communicating degrees of satisfaction. Kaye et al. (2017) go as far as suggesting them for consideration as possible alternatives to regular Likert scales.

Emoji Semantics. The MIT DeepMoji project (Felbo et al., 2017) developed a model that recommends emojis given a natural language sentence as input. A deep neural architecture was trained on a collection of 1.2B tweets to learn the sentiment, emotions, and the use of sarcasm in short text.

Barbieri et al. (2016b) proposed a method to learn vector space embeddings of emojis using the standard word2vec skip-gram approach, applied to a large collection of tweets. In contrast, Eisner et al. (2016) attempted to learn vector embeddings of emojis based on their short descriptions in the Unicode standard. EmojiNet (Wijeratne et al., 2017) provides a sense inventory to distinguish different senses of an emoji, drawing on Web-crawled emoji definitions and connecting them to word senses from a lexical resource, along with vector representations of context words.

The first paper to thoroughly investigate the sentiment of emojis (Novak et al., 2015) proposed a sentiment ranking of 715 emojis on a corpus of 70,000 tweets. This work provides a basis for future research on the logographic usage of emojis in social media. Rakhmetullina et al. (2018) classify 15 emojis with regard to their sentiment polarity and with regard to 4 emotion classes. For this, they applied a distant supervision technique for a reliable mapping based on manually annotated data. Li et al. (2019) used a heuristic to observe ties between emojis and emotions in social media data and compared emoji usage on Twitter and Weibo. Their heuristic involves training word vector models and then invoking a word–emotion lexicon to obtain average vectors for 8 emotions. Finally, EmoTag (Shoeb et al., 2019) provides interpretable word vectors that describe words in terms of their association with emojis. These vectors were found to be useful for emotion prediction.

Zhou and Wang (2017) trained a natural language conversation model that accounts for the underlying emotion of utterances by exploiting the existence of emojis as a signal.

3 Annotation Task

In order to better study the connection between emojis and emotions, we proceeded to compile a dataset of ratings quantifying the perceived strength of association between emojis and emotions.

3.1 Task Setup and Guidelines

Target Emoji Set. We considered a set of 150 most frequently used emojis, based on frequencies reported by the Emoji Tracker service[2], a platform that visualizes the real-time use of emojis on Twitter. The counters on Emoji Tracker indicate how many times an emoji has been used on Twitter since July 4, 2013. We rank all emojis based on their reported total frequency counts as of July 3, 2019 and pick the top 150 emojis for our annotation task. While their frequencies are based on global data, the ranking remains useful because of the large proportion of English tweets (Vicinitas, 2018) and the fact that emoji use is broadly similar across languages (Barbieri et al., 2016a), despite certain language-specific differences.

Emotion Set. While numerous emotion models and affective classification schemes have been put forth, for this study we consider the 8 basic emotions proposed in the Wheel of Emotions model by Plutchik (1980), i.e., *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, and *trust*.

Linguistic Context. In this study, we focus on emoji use within the English language. A previous study found that the meaning of an emoji remains relatively stable across different languages and media (Barbieri et al., 2016a). In part, this may stem from the language-independent visual nature of emojis. However, different concepts may have different associations in different cultures, so our results cannot be taken as being universal.

Ratings. For a given emoji, the participants were asked to assess to what extent said emoji is associated with a given emotion, for each of the 8 different target emotions.

Association is a broad notion that not only covers emojis that are directly invoked to express an emotion, as in the case of certain facial expression emojis, but also encompasses mere conceptual association. For instance, the wrapped gift emoji may be associated with *joy*, although the semantics of the emoji itself correspond to a *present* or *gift* rather than directly conveying *joy*.

Note also that this notion of association reflects a general, abstract form of connection, much like a prior. Clearly, embedded in a specific utterance, the specific emotions that are evoked may differ quite substantially, due to the complex ways in which different words along with embedded emojis interact to give rise to an overall interpretation. In this regard, our ratings are similar to widely used word relatedness resources that seek to quantify context-independent lexical associations (Finkelstein et al., 2001) or word–emotion associations (Mohammad and Turney, 2013; Mohammad, 2018).

The degree of association was specified numerically as a score ranging from 0 (no association with the emotion) to 4 (representing the highest degree of association with the emotion). While we are cognizant of the challenges of directly eliciting scalar ratings from the annotators, we opted to follow prominent previous work on collecting association ratings (Rubenstein and Goodenough, 1965; Finkelstein et al., 2001; Hill et al., 2015; Gerz et al., 2016) in order to make our data comparable to such efforts.

3.2 EmoTag1200 Data Collection

Interface. We developed a web interface to collect ratings. We randomly split the target set of 150

[2] http://emojitracker.com/

emojis into a total of 6 subsets, each consisting of 25 emojis. When a rater selects a set from the main page, the corresponding 25 emojis are presented to the user alongside their official names, each to be annotated with respect to our set of 8 different emotions.

Within each set, we randomize the order of displayed emojis upon each page load, such that different raters do not observe and annotate them in the same order. This ensures that different emojis within a set are given equal attention on average when aggregating scores from different human raters, mitigating potential fatigue-driven biases in the final ratings.

In total, an annotator makes 8 selections for a single emoji, corresponding to the set of 8 basic emotions. We ask users to provide all 8 emotions ratings for every single emoji. This is because a single emoji may be tied to different kinds of emotions. For example, the Kiss Mark emoji may express joy, trust, anticipation, among others. This is why our annotation task was designed to solicit scores ranging from 0 to 4 for eight different emotions for each individual emoji.

Participants. We recruited a total of 9 different human participants to each rate 150 emojis for 8 different emotions. All selected participants were from the age group between 25 and 35 years and native or near-native speakers of English who reported having extensive prior familiarity with emojis in their personal communication or from social media use. As mentioned, the emojis were grouped into 6 sets, each consisting of 25 emojis. The annotators were asked to annotate one such set per day so as to avoid overburdening them, which might affect the quality of the rating.

The original intensity scores range from 0 to 4, but are rescaled to $[0, 1]$. Ultimately, for each pairing of emoji and emotion, we consider the mean value across the 9 individual raters as a real-valued score in $[0, 1]$ reflecting the association for that pairing. We also compute for each pairing the standard deviation among its ratings.

3.3 Analysis

In total, we collect 10,800 ratings for 1,200 pairings of emoji and emotion, covering 150 emojis, each rated with regard to 8 emotions by 9 human raters.

Inter-Annotator-Agreement. To evaluate the agreement between the raters, we first check the overall agreement between pairs of human raters across the entire set of emoji–emotion ratings. This was in part also motivated by quality control concerns, i.e., a desire to assess whether there was any individual rater that disagreed substantially with all other raters. Fortunately, this was not the case and we decided not to eliminate data from any rater. Figure 1 reports the pairwise Pearson correlation scores between raters.

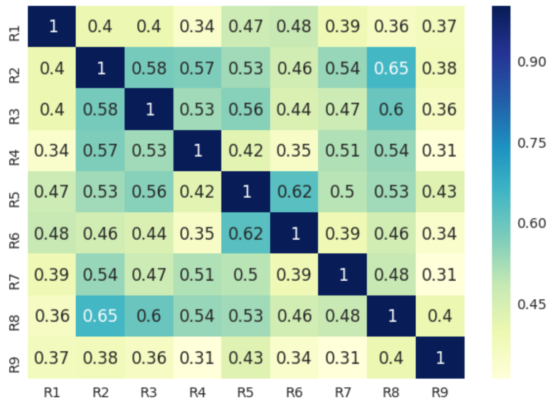

Figure 1: Pairwise Pearson correlation for 9 raters based on all 8 emotion scores for the set of 150 most popular Twitter emojis

We focus on Pearson correlation in this analysis in order to later be able to compare these scores against Pearson correlation scores obtained when comparing automated prediction methods against the ground truth (Section 4). In Figure 2, we consider separately for different emotions the average agreement (Pearson correlation) of raters with the mean ratings. We find that a fairly high agreement is observed for *sadness*, *joy*, and *fear*. In contrast, we conjecture that for *surprise*, *trust*, and *anticipation*, it appears somewhat less obvious which emojis one would normally use to convey such emotions. Instead, we observe that individual annotators sometimes provided high rating scores based on idiosyncratic associations. One rater, for instance, associated a gemstone with a high degree of anticipation, while the others did not. It is important to be aware of these varying correlation scores and compute separate correlation scores per emotion when evaluating emotion prediction models on this data. In Figure 3, we visualize the emotion-specific agreement for different individual raters.

Emoji-Specific Agreement. We also invoke Krippendorff's α as a measure of agreement between raters for each individual emoji along with its emotions. This allows us to understand to what

Emoji	Name	Emotion	All Ratings	Emotion Score	K. α	SD σ
U+1F621 😡	Pouting Face	Anger	1.00 (9×)	1.00	0.61	0.00
U+1F60A 😊	Smiling Face with Smiling Eyes	Joy	0.75 (3×), 1.00 (6×)	0.92	0.68	0.12
U+1F62D 😭	Loudly Crying Face	Sadness	1.00 (9×)	1.00	0.48	0.00
U+1F633 😳	Flushed Face	Fear	0.00 (2×), 0.25 (2 ×), 0.50, 0.75 (2×), 1.00 (2×)	0.50	0.12	0.37
U+1F449 👉	Backhand Index Pointing Right	Anticipation	0.00 (6×), 0.5, 0.75, 1.00	0.25	0.11	0.37

Table 1: Examples emoji emotion ratings along with Krippendorff's (K) α and Standard Deviation (SD) σ.

Groups	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
B4 ($\geq$ 0.75)	3	1	0	3	**23**	6	2	2
B3 ($\geq$ 0.50)	3	5	14	8	**24**	8	5	**24**
B2 ($\geq$ 0.25)	19	**86**	20	18	35	13	33	38
B1 ($\geq$ 0.00)	**125**	58	116	121	68	123	110	86

Table 2: The distribution of 150 target emojis across four buckets B1, B2, B3, and B4 with respect to their gold intensity score for all 8 emotions. The bold score represents which emotion gets the highest number of emojis in the respective bucket.

Emotions	Top 150 Emojis (EmoTag1200)	Other Predicted Emojis (excluding Top 150)
Anger	😠 Angry Face	👺 Japanese Goblin
	😡 Pouting Face	👹 Japanese Ogre
	😤 Face with Steam from Nose	🔪 Hocho
Anticipation	👀 Eyes	🎆 Fireworks
	💭 Thought Balloon	🌠 Shooting Star
	💰 Money Bag	👰 Person with Veil
Disgust	😖 Confounded Face	🐜 Ant
	😣 Persevering Face	👹 Japanese Ogre
	👎 Thumbs Down	😲 Astonished Face
Fear	😨 Fearful Face	🔪 Hocho
	😱 Face Screaming in Fear	👹 Japanese Ogre
	😰 Anxious Face with Fear	👺 Japanese Goblin
Joy	☺️ Smiling Face	🎂 Birthday Cake
	😆 Grinning Squinting Face	🎊 Confetti Ball
	😂 Face with Tears of Joy	💝 Heart with Ribbon
Sadness	😢 Crying Face	🔪 Hocho
	😭 Loudly Crying Face	👼 Baby Angel
	💔 Broken Heart	😿 Crying Cat Face
Surprise	‼️ Double Exclamation Mark	😮 Face with Open Mouth
	❗ Exclamation Mark	😵 Dizzy Face
	😱 Face Screaming in Fear	🙀 Weary Cat
Trust	😙 Kissing Face with Smiling Eyes	⛪ Church
	💕 Two Hearts	👪 Family
	🌹 Rose	⚓ Anchor

Table 3: Three top-ranked emotion-intensive emojis for eight emotions, considering ground truth annotations from the top 150 emojis in our ground-truth EmoTag1200 dataset on the left, and using unsupervised emotion intensity predictions for the remaining emojis (i.e., those not included in the set of 150) on the right.

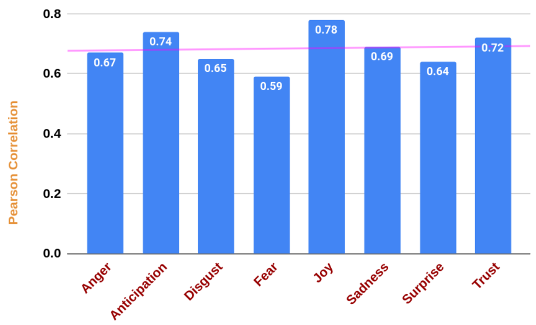

Figure 2: Average Pearson correlation coefficient between rater score and the gold score grouped by emotions. The pink line represents the overall trend.

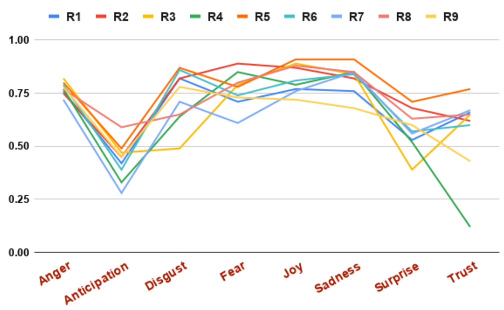

Figure 3: Variation of Pearson correlation coefficient for individual raters with the mean score across emotions

extent the raters agree or disagree on the rating of a *particular emoji–emotion* pairing. The scores range from 0 to 1, where $\alpha = 0$ denotes no agreement and $\alpha = 1$ represents the highest level of agreement among all users. Table 1 shows a few examples of emojis with specific emotions and their associated ratings, including the Krippendorff α value and standard deviation. We include examples with high as well as low agreement.

Distribution and Examples. In Table 2, we report the distribution of scores for different emotions. As one might reasonably expect, the lowest-intensity bucket is the largest for each considered emotion. Overall, fairly few emojis are strongly associated with *anger*, *disgust*, *fear*, *sadness*, or *surprise*. For *disgust*, no emoji falls into the highest-intensity bucket, although some show a moderate intensity level. There are numerous emojis associated with *anticipation*. The most atypical distribution is observed for *joy*, as there appear to be a wide range of objects and concepts that spark joy, in addition to the emojis that directly express joy.

Finally, in Table 3, we list the top-ranked 3 emotion-bearing emojis for each of the 8 considered emotions based on our dataset ("Top 150 Emojis" column) as well as based on an automated prediction for other emojis not in our annotated dataset (described later in Section 4.3). Indeed, for many emotions, we encounter some of the most prototypically expected emojis, especially facial expression ones. Note that in some cases, common use diverges from the original Unicode definitions of the emojis, as for instance for the "Persevering Face" 😣 emoji, which is also associated with *disgust* rather than just with perseverance.

4 Emotion Scoring Experiments

Given our manually collected data for 150 emojis, we next consider to what extent simple *unsupervised* methods and resources correlate with these associations such that they could be used to reproduce such associations automatically in a data-driven manner. The EmoTag1200 data compiled in Section 3, specifically the mean ratings for emoji–emotion pairs, serve as the ground truth.

4.1 Corpus Data

To enable an unsupervised prediction, we explore methods relying on several different kinds of resources, including existing pre-trained word embedding models and word emotion lexicons, which will be described later on in Section 4.2 when introducing the specific methods.

Additionally, we make use of distributional similarity to support several of the methods. For this, we draw on an emoji-centric corpus. In order to infer the correlation of emojis with emotion-bearing words and vice versa, we created a web crawl of tweets collected specifically to provide emoji statistics by seeking out tweets containing at least one emoji. We consider a set of 620 most frequently used emojis from Novak et al. (2015) and from Emoji Tracker. For each emoji, we then retrieved an equal number of tweets labeled as being in English. In total, we obtained a set of 20.8 million tweets over a span of one year (Shoeb et al., 2019).

Subsequently, we train simple 300-dimensional word2vec skip-gram (Mikolov et al., 2013) models on this corpus. As this corpus contains numerous occurrences of emojis, the resulting word vector representations include vectors for emojis, and we are able to compute the cosine similarity between emojis and words.

In the following, we explain how this data comes into play while predicting emotion ratings for any emojis available in our corpus.

4.2 Prediction Methods

We consider several methods to predict emoji–emotion association scores. These include methods that directly consult distributed word vectors, as well as methods that draw on different kinds of word emotion lexicons.

Distributed Word Vectors based on Emotion Words. The first method we consider is to directly rely on standard distributed word embeddings E (with vocabulary $\mathcal{V}_E$), as these have been shown to carry emotional associations (Raji and de Melo, 2020). Given an emoji e and an emotion (affect) a, we consult E attempting to obtain a vector $\mathbf{v}_e$ for the emoji as well as a vector $\mathbf{v}_{l_a}$ for the word l_a that serves as a label for the affect a (e.g., the words *joy*, *anger*, etc.). We then compute the association in terms of the cosine similarity and treat it as the rating:

$$\sigma(e,a) = \begin{cases} \mathrm{sim}(\mathbf{v}_e, \mathbf{v}_{l_a}) & e, l_a \in \mathcal{V}_E \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Here, $\mathrm{sim}(\mathbf{v}_1, \mathbf{v}_2)$ denotes the cosine similarity between two vectors.

We first consider the widely used 300-dimensional GloVe (Pennington et al., 2014) models pretrained on CommonCrawl 840B and Twitter, as these contain emojis. However, given that their emoji coverage is limited, we additionally consider word2vec (Mikolov et al., 2013) skip-gram models that we trained on our crawled Twitter data from Section 4.1, using window sizes of 5 and 25.

Binary Word Emotion Lexicons based on Emoji Corpus Similarities. We next consider a series of approaches that rely on word emotion lexicons in conjunction with our emoji corpus to connect these lexicons to emojis. EmoLex by Mohammad and Turney (2013), also known as the NRC Emotion Lexicon, is among the most prominent English language word emotion lexicons. It assigns words binary labels for the same eight emotions that we consider in our study. Thus, a word may either be tagged as being associated with *trust* or as not being associated with it. Specifically, we consult EmoLex to find the subset of words $\mathcal{V}_a$ from the vocabulary $\mathcal{V}$ that are associated with affect a.

To find a connection between emojis and words in the lexicon, we again draw on our emoji corpus from Section 4.1. We rank the top $k = 5$ words from the lexicon's $\mathcal{V}_a$ based on the word–emoji cosine similarities induced from our corpus, and finally compute an emoji e's emotion score $\sigma(e,a)$ for affect a as the average of similarity scores for the top k words:

$$\sigma(e,a) = \frac{1}{k} \sum_{w \in \mathrm{T}(e,\mathcal{V}_a)} \mathrm{sim}(\mathbf{v}_e, \mathbf{v}_w) \quad (2)$$

Here, $\mathbf{v}_e$ denotes our emoji corpus vector embedding for an emoji e, while $\mathbf{v}_w$ denotes our emoji corpus vector embedding for a word w. Such words are taken from $\mathrm{T}(e,\mathcal{V}_a)$, defined as the set of top-k words w in $\mathcal{V}_a$, i.e., among those words tagged as having the affect a in the lexicon, ranked in terms of $\mathrm{sim}(\mathbf{v}_w, \mathbf{v}_e)$ scores, where $\mathrm{sim}(\mathbf{v}_1, \mathbf{v}_2)$ again denotes the cosine similarity between two vectors. Note that a top-k word can be considered only if it is available in the binary word emotion lexicon (EmoLex). Indeed, some potentially valuable out-of-vocabulary (OOV) word forms are disregarded, as they do not have any available emotion labels. Examples of such top-ranked OOV word forms are *helooooo*, *funnnn*, etc.

Word Emotion Intensity Lexicons using Emoji Corpus Similarities. Next, we consider emotion lexicons that, unlike EmoLex, provide real-valued emotion scores for English words. In this case, the emotion intensity scores of words directly figure into the predicted scores. We first consult the lexicon to find all words $\mathcal{V}_a$ for which the lexicon provides any emotion intensity score at all for affect a. We then identify the top k words in terms of the word–emoji cosine similarity scores based on our emoji corpus, as earlier. Finally, however, our predicted score $\sigma(e,a)$ is the arithmetic mean of emotion intensity scores of the top k words. Specifically,

$$\sigma(e,a) = \frac{1}{k} \sum_{w \in \mathrm{T}(e,\mathcal{V}_a)} \tau(w,a), \quad (3)$$

where $\tau(w,e)$ denotes the emotion intensity score provided by the lexicon and the remaining variables are defined as earlier.

In our experiments with this approach, we consider two separate word emotion lexicons: the NRC Emotion Intensity lexicon (NRC-EIL) by Mohammad (2018) and DepecheMood++ by Araque et al.

Source	Variant	Anger	Anticip.	Disgust	Fear	Joy	Sadness	Surprise	Trust	Average
Distributed Word Vectors via Emotion Words										
GloVe	CommonCrawl	0.05	0.05	0.11	0.09	0.34	-0.05	**0.18**	0.18	0.12
	Twitter	0.08	0.02	0.07	-0.04	0.07	-0.11	0.06	0.06	0.03
word2vec	Window=5	0.60	0.15	0.74	0.48	0.50	0.50	0.32	-0.16	0.39
(Emoji Corpus)	Window=25	0.64	0.12	0.69	0.57	0.63	0.56	0.42	-0.05	0.45
Emotion Lexicons via Emoji Corpus Word Similarities										
EmoLex	k =5	0.62	-0.03	**0.81**	0.50	0.19	0.57	-0.27	-0.04	0.29
NRC-EIL	k =10	0.35	0.24	0.23	0.44	0.60	0.50	0.05	0.42	0.35
	k =100	0.57	0.21	0.62	0.71	0.71	0.74	0.06	0.46	0.51
	k =300	0.60	0.21	0.69	**0.72**	0.71	0.71	0.03	**0.50**	0.52
DepecheMood++	k =10	0.32	N/A	N/A	0.54	0.32	0.5	0.09	N/A	0.35
	k =100	0.43	N/A	N/A	0.61	0.35	0.72	0.08	N/A	0.41
	k =300	0.48	N/A	N/A	0.69	0.37	0.76	0.05	N/A	0.44
Emotion Lexicons via Emoji Corpus Co-Occurrence Frequencies										
NRC-EIL	k =300	**0.74**	**0.23**	0.59	0.65	**0.72**	0.74	0.08	0.46	**0.53**
DepecheMood++	k =200	0.58	N/A	N/A	0.37	0.41	**0.78**	0.10	N/A	0.44
Human Annotation										
Human Agreement		0.67	0.74	0.65	0.59	0.78	0.69	0.64	0.72	0.69

Table 4: Pearson Correlation scores for all considered prediction methods. Bolded scores represent the highest correlation observed for the emotion in the respective column except for the human agreement score.

(2018). The latter has a different emotion inventory than the Plutchik (1980) emotion labels that we rely upon, so we apply the following mapping: *angry* $\mapsto$ *anger*, *afraid* $\mapsto$ *fear*, *happy* $\mapsto$ *joy*, *sad* $\mapsto$ *sadness*, and *amused* $\mapsto$ *surprise*. DepecheMood++ is an automatically constructed lexicon that provides frequencies of each word along with their emotion score. We apply a minimal frequency threshold of 50, as this was found to eliminate less reliable entries.

Word Emotion Intensity Lexicons using Emoji Corpus co-occurrences. Finally, we further consider a variant of the above formula, where $\mathrm{T}(e, \mathcal{V}_a)$ does not rank words in terms of word2vec cosine similarities, but instead based on their co-occurrence frequency with the emoji e in our Twitter corpus.

4.3 Results

Table 4 compares the mean human-annotated emotion ratings from EmoTag1200 against predicted scores induced using the aforementioned methods, evaluated in terms of Pearson correlation coefficients.

The pretrained GloVe embeddings exhibit very low correlations, as both models have a limited coverage of just 26 out of the 150 emojis in the ground truth data. Our emoji-centric corpus yields stronger results. Among the two variants, word vectors trained with a larger context window size of 25 perform better, because emojis are often placed at the end of tweets. This result also accords with previous studies that show that larger context windows tend to capture generic relatedness, while shorter ones emphasize functional similarity of words (Levy and Goldberg, 2014).

Using EmoLex with our binary emotion label scores, we observe varied results, including strong correlation for *disgust*, but low or even negative for several others. This is because the EmoLex lexicon merely signals whether or not it considers a word as being associated with an emotion. Such binary emotion labels do not appear to convey sufficient information for a more accurate prediction.

With the NRC Emotion Intensity lexicon (Mohammad, 2018), we are able to obtain substantially higher correlations for a range of different settings of top-k words, both with our emoji corpus vector similarity as well as with co-occurrence frequency rankings. Thus, high-quality emotion lexicons providing crowdsourced emotion intensity ratings provide valuable information beyond what distributed word vectors deliver directly.

DepecheMood++, owing to its automatic data-driven induction process, does not yield as good results as the high-quality crowdsourced scores compiled in the NRC Emotion Intensity lexicon. Moreover, DepecheMood++ does not cover all emotions in the ground truth dataset.

Overall, we find that we are able to obtain a high correlation with the human ratings in EmoTag1200. Thus, we apply our models to predict scores for a larger set of 620 emojis from our emoji corpus. In

Table 3, we list the top 3 emojis for each emotion in terms of the predictions using the similarity-based approach with NRC-EIL (k=300), but excluding any emojis already in our EmoTag1200 ground truth data. The results (column labeled "Other Predicted Emojis") show that we are automatically able to find additional emojis tied to emotions.

5 Conclusion

The desire to express an emotion is one of the factors that has driven the tremendous proliferation of emojis in interpersonal communication. However, this connection has not been studied in sufficient detail, at the level of individual emojis. In this work, we shed light on this connection by compiling the EmoTag1200 dataset, which quantifies people's reported association between emojis and emotion. From each of 9 human raters, we solicit 1,200 ratings covering a set of 150 emojis with regard to 8 core emotions from Plutchik (1980)'s Wheel of Emotions. This constitutes the first resource of this kind, which we thoroughly analyze and make freely available to enable further research.

An important avenue of future work will be to assess to what extent there may be cultural differences in these associations (see Discussion in Section 3.1). Similarly, variation with respect to age and other variables merits further study as well. Temporal aspects could be considered in diachronic studies, to account for the fact that emoji use has been evolving.

Finally, we rely on our annotated data to study how well we can automatically estimate emotional association ratings for a given emoji, considering a series of different baseline methods and resources. Our findings suggest that data-driven methods can fare quite well at this if combined with high-quality affective intensity information at the lexical level. Hence, we are able to predict high-quality emotion scores for a larger set of emojis.

This opens up further research avenues on possible downstream applications exploiting this knowledge. The most obvious use cases are sentiment analysis (Dong and de Melo, 2018), emotion analysis (Raji and de Melo, 2020), consumer behaviour analytics (Dong et al., 2020), context-sensitive emoji recommendation (Felbo et al., 2017), computational social science and public opinion mining (Wang et al., 2018; Du et al., 2020), and user modeling (Guo et al., 2018), but it may also be useful in dialogue systems (Delobelle and Berendt, 2019), e.g. to detect sarcasm. As emoji use is now ubiquitous on mobile devices and social media, we believe that ultimately any NLP task involving social media text may benefit from such emoji resources.

Acknowledgments

We sincerely thank all our annotators who volunteered to contribute to our study and helped us to establish this new resource for the community. Clearly, this research would not have been possible without their time and effort.

References

Oscar Araque, Lorenzo Gatti, Jacopo Staiano, and Marco Guerini. 2018. DepecheMood++: a bilingual emotion lexicon built through simple yet powerful techniques. *arXiv preprint arXiv:1810.03660*.

Francesco Barbieri, Miguel Ballesteros, and Horacio Saggion. 2017. Are emojis predictable? In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 105–111, Valencia, Spain. Association for Computational Linguistics.

Francesco Barbieri, German Kruszewski, Francesco Ronzano, and Horacio Saggion. 2016a. How cosmopolitan are emojis?: Exploring emojis usage and meaning over different languages with distributional semantics. In *Proceedings of the 24th ACM International Conference on Multimedia*, MM '16, pages 531–535, New York, NY, USA. ACM.

Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. 2016b. What does this emoji mean? a vector space skip-gram model for Twitter emojis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3967–3972, Portorož, Slovenia. European Language Resources Association (ELRA).

Lyn Bartram, Abhisekh Patra, and Maureen Stone. 2017. Affective color in visualization. In *Proceedings of CHI 2017*, pages 1364–1374. ACM.

Prashant Bordia. 1997. Face-to-face versus computer-mediated communication: A synthesis of the experimental literature. *The Journal of Business Communication (1973)*, 34(1):99–118.

Spencer Cappallo, Stacey Svetlichnaya, Pierre Garrigues, Thomas Mensink, and Cees G. M. Snoek. 2019. New modality: Emoji challenges in prediction, anticipation, and retrieval. *IEEE Transactions on Multimedia*, 21(2):402–415.

Charles Darwin. 1872. *The Expression of the Emotions in Man and Animals*. Appleton. The original was published 1898 by Appleton, New York. Reprinted 1965 by the University of Chicago Press, Chicago and London.

Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2010. Enhanced sentiment learning using Twitter hashtags and smileys. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, COLING '10, pages 241–249, Stroudsburg, PA, USA. Association for Computational Linguistics.

Pieter Delobelle and Bettina Berendt. 2019. Time to take emoji seriously: They vastly improve casual conversational models. *arXiv preprint arXiv:1910.13793*.

Daantje Derks, Agneta H. Fischer, and Arjan E.R. Bos. 2008. The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior*, 24(3):766 – 785. Instructional Support for Enhancing Students' Information Problem Solving Ability.

Xin Dong and Gerard de Melo. 2018. A helping hand: Transfer learning for deep sentiment analysis. In *Proceedings of ACL 2018*, pages 2524–2534.

Xin Dong, Jingchao Ni, Wei Cheng, Zhengzhang Chen, Bo Zong, Dongjin Song, Yanchi Liu, Haifeng Chen, and Gerard de Melo. 2020. Asymmetrical hierarchical networks with attentive interactions for interpretable review-based recommendation. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020)*. AAAI Press.

Xu Du, Matthew Kowalski, Aparna Varde, Gerard de Melo, and Robert W. Taylor. 2020. Public opinion matters: Mining social media text for environmental management. *ACM SIGWEB Newsletter*, Autumn 2019.

Ben Eisner, Tim Rocktäschel, Isabelle Augenstein, Matko Bosnjak, and Sebastian Riedel. 2016. emoji2vec: Learning emoji representations from their description. In *Proceedings of The Fourth International Workshop on Natural Language Processing for Social Media*, pages 48–54, Austin, TX, USA. Association for Computational Linguistics.

Paul Ekman and Wallace V. Friesen. 1986. A new pan-cultural facial expression of emotion. *Motivation and Emotion*, 10(2):159–168.

Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1615–1625, Copenhagen, Denmark. Association for Computational Linguistics.

Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, and Eytan Ruppin. 2001. Placing search in context: The concept revisited. In *Proceedings of the 10th International Conference on World Wide Web*, WWW '01, pages 406–414, New York, NY, USA. ACM.

Daniela Gerz, Ivan Vulić, Felix Hill, Roi Reichart, and Anna Korhonen. 2016. SimVerb-3500: A large-scale evaluation set of verb similarity. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2173–2182, Austin, Texas. Association for Computational Linguistics.

Alec Go, Richa Bhayani, and Lei Huang. 2009. Twitter sentiment classification using distant supervision. Technical report. Stanford University.

Ziyu Guo, Liqiang Wang, Yafang Wang, Guohua Zeng, Shijun Liu, and Gerard de Melo. 2018. Public opinion spamming: A model for content and users on sina weibo. In *Proceedings of the 10th ACM Conference on Web Science*, pages 210–214. ACM.

Felix Hill, Roi Reichart, and Anna Korhonen. 2015. Simlex-999: Evaluating semantic models with (genuine) similarity estimation. *Computational Linguistics*, 41(4):665–695.

Linda K. Kaye, Stephanie A. Malone, and Helen J. Wall. 2017. Emojis: Insights, affordances, and possibilities for psychological science. *Trends in Cognitive Sciences*, 21(2):66 – 68.

Tugba Kulahcioglu and Gerard de Melo. 2018. FontLex: A typographical lexicon based on affective associations. In *Proceedings of the 11th Language Resources and Evaluation Conference (LREC 2018)*, Paris, France. European Language Resources Association (ELRA).

Tugba Kulahcioglu and Gerard de Melo. 2019. Paralinguistic recommendations for affective word clouds. In *Proceedings of ACM IUI 2019*, pages 132–143, New York, NY, USA. ACM.

Tugba Kulahcioglu and Gerard de Melo. 2020. Fonts like this but happier: A new way to discover fonts. In *Proceedings of ACM Multimedia 2020*, New York, NY, USA. ACM.

Peter J. Lang, Margaret M. Bradley, and Bruce N. Cuthbert. 1999. International affective picture system (IAPS): Technical manual and affective ratings. Technical report, University of Florida.

Omer Levy and Yoav Goldberg. 2014. Dependency-based word embeddings. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 302–308, Baltimore, Maryland. Association for Computational Linguistics.

Mingyang Li, Sharath Guntuku, Vinit Jakhetiya, and Lyle Ungar. 2019. Exploring (dis-)similarities in emoji-emotion association on Twitter and Weibo. In *Companion Proceedings of The 2019 World Wide Web Conference*, WWW '19, page 461–467, New York, NY, USA. Association for Computing Machinery.

Gretchen McCulloch. 2019. *Because Internet: Understanding the New Rules of Language*. Penguin Publishing Group.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc.

Saif M. Mohammad. 2018. Word affect intensities. In *Proceedings of the 11th Language Resources and Evaluation Conference*, Miyazaki, Japan. European Language Resource Association.

Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465.

Noa Na'aman, Hannah Provenza, and Orion Montoya. 2017. MojiSem: Varying linguistic purposes of emoji in (Twitter) context. In *Proceedings of ACL 2017, Student Research Workshop*. Association for Computational Linguistics.

Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. 2015. Sentiment of emojis. *Plos One*, 10(12).

Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global vectors for word representation. In *EMNLP*.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In Robert Plutchik and Henry Kellerman, editors, *Theories of Emotion*, pages 3–33. Elsevier.

Shahab Raji and Gerard de Melo. 2020. What sparks joy: The AffectVec emotion database. In *Proceedings of The Web Conference 2020*, pages 2991–2997, New York, NY, USA. ACM.

Aisulu Rakhmetullina, Dietrich Trautmann, and Georg Groh. 2018. Distant supervision for emotion classification task using emoji 2 emotion. In *Proceedings of the 1st International Workshop on Emoji Understanding and Applications in Social Media (Emoji2018)*, Stanford, CA, USA.

Herbert Rubenstein and John B. Goodenough. 1965. Contextual correlates of synonymy. *Commun. ACM*, 8(10):627–633.

Abu Awal Md Shoeb, Shahab Raji, and Gerard de Melo. 2019. EmoTag – Towards an emotion-based analysis of emojis. In *Proceedings of RANLP 2019*, pages 1094–1103.

Vicinitas. 2018. 2018 research on 100 million tweets: What it means for your social media strategy for Twitter. Online: https://www.vicinitas.io/blog/twitter-social-media-strategy-2018-research-100-million-tweets.

Leticia Vidal, Gastón Ares, and Sara R. Jaeger. 2016. Use of emoticon and emoji in tweets for food-related emotional expression. *Food Quality and Preference*, 49:119–128.

Tamara Voronova and Andrei Sterligov. 1997. *Western European Illuminated Manuscripts of the 8th to the 16th centuries*. Parkstone Press.

Liqiang Wang, Ziyu Guo, Yafang Wang, Zeyuan Cui, Shijun Liu, and Gerard de Melo. 2018. Social media vs. news media: Analyzing real-world events from different perspectives. In *Proceedings of DEXA 2018*, volume 11030 of *LNCS*, pages 471–479. Springer Verlag.

Sanjaya Wijeratne, Lakshika Balasuriya, Amit P. Sheth, and Derek Doran. 2017. EmojiNet: An open service and API for emoji sense discovery. *CoRR*, abs/1707.04652.

Rui Zhou, Jasmine Hentschel, and Neha Kumar. 2017. Goodbye text, hello emoji: Mobile communication on WeChat in China. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 748–759, New York, NY, USA. Association for Computing Machinery.

Xianda Zhou and William Yang Wang. 2017. MojiTalk: Generating emotional responses at scale. *arXiv 1711.04090*.