

# How reliable are emojis in representation and generation of emotions?

Meghdut Sengupta  
Ethics and NLP  
University of Stuttgart  
29.09.2020

## Abstract

It's interesting to observe the evolution of social media over the past few decades and the part played by emojis in expressing emotions in that process. In recent times emojis have been used as a modality for numerous tasks in sentiment analysis. However there are a few questions about reliability, security and privacy concerns regarding the usage of emojis in the modern day era. This report is aimed at presenting three sources concerning interdisciplinary studies regarding emojis, as far as machine learning is concerned. Firstly, I shall describe an article by Jun Wu(7) which talks about incorporation of empathy in Artificial General Intelligence (AGI)(7) and how it will be helpful in making AI more ethical. Secondly, I'll take a look into emotion generation with the help of emojis from the paper 'Mojitalk: Generating Emotional Responses at Scale' by Zhou et al.(8) Thirdly, 'The not-so-cute side of emojis: Potential security, privacy, and bandwidth issues' - an opinion on emoji usage in the modern world by Jack Wallen(6), will be discussed. The summary of the discussion session of my presentation has also been included at the very end.

## 1 Introduction

Since an AI system would need to possess transparency in order to develop interactions with humans for its betterment, inducing emotional intelligence in AI is gaining popularity. One of the prime use cases in this field is robots venturing in teaching autistic children.(1) However to really understand the emotions that human beings experience, an AI must also have the same persistent capacity to realize and act upon emotion generation accordingly. For example, the emotion chip of EPU developed by Emoshape can understand 64 trillion possible emotional states every 1/10th of a second. The range of emotions is mapped onto a

gradient where the degree of each emotion can be observed.(7) In the modern trends, one very important way to generate emotions is basically by leveraging the usage of emojis. While emojis are one of the primary means of expressing emotions, there are a lot of scopes for exploration in this domain. Zhou et al uses some number of twitter data to associate emoticons to a target document. Obviously, a concern remains as to how this process can be made ethically more secure. As per Jack Wallen, in his article 'The not-so-cute side of emojis: Potential security, privacy, and bandwidth issues', usage of emojis in business communication can potentially lead to complicated situations. More will be discussed on the privacy aspects in the section of 'Emojis - Privacy and Security'.

## 2 Empathy in AGI

In the broader context empathy is in Self-Awareness, Social-Awareness, Self-Actualization and Transcendence.(7) AI startup advisor Steve Ardire says "People don't change behavior on information, they change it on emotion, emotional intelligence, and empathy."(7) In other words, its not possible to have 'human-like' intelligence without personality or emotions. Achieving emotional intelligence is roughly a three step process. They are defined as Cognitive empathy, Affective empathy and Somatic empathy.(7)

Cognitive empathy is the ability to see a certain situation from the perspective of the person who's facing it. In other words its about stepping into the shoes of the person and viewing the situation from his perspective. Affective empathy corresponds to the ability to understand and respond to the situation. Somatic empathy is more like the effect of the Affective empathy, where its characterized by the physical response of the event. These three factors are similar in both human and artificial em-

pathy on a generic level.

But on a deeper level, just like human-intelligence is different from artificial intelligence, artificial empathy is also different from human empathy. "Heartificial Empathy"(2) - by Minster Dial describes artificial empathy as programming empathy in machines. Based on this difference, the very popular question surfaces on the ethical grounds of AI. "Is it really for the best of humankind?"

In her article, Wu says that, since our fear for singularity is justified, "Uncertainty just means that we need to explore the business cases for empathy in our AI Systems". She argues that for incorporation of AI in specific business cases can "drastically" increase the prospects of ethical business practices. In her opinion, better ethical practices in business practices amounts to higher ethical standards of living.

One example cited by her is that a common problem in the healthcare sector is that the doctors and the nurses often report a 'burn-out' in their work. Wu argues that ethically trained AI robots can use sufficient levels of empathy to work in places of human health professionals. Hence they can divide the workload to relieve the humans of the stress they face for overworking.

However I do not feel very convinced with either the claim that better ethical standards in business will improve our lifestyle or the example cited to back it. In the modern world of crony capitalism its really hard to think that if, at all, we build a successful AI which can empathize perfectly with human beings, then it will be used to 'divide' the workload between humans and not rather 'replace' them. In the latter case, which seems to be more likely to me, a large section of the human population will be left unemployed. It feels like she explores and presents an ideal case with an assumption that we have a hundred percent control over the AI we are coding, while in fact, explainable AI says that we, in fact, have little understanding regarding what happens inside the black box of the models in deep neural network.(5)

The second claim made by her is that empathetic AI can make better decisions. She said when an active chatbot detects happiness in the customer at the end of the conversation, then it has the true confirmation of satisfaction from the customer. So the experience of talking to a chatbot in that case will be more "personal".

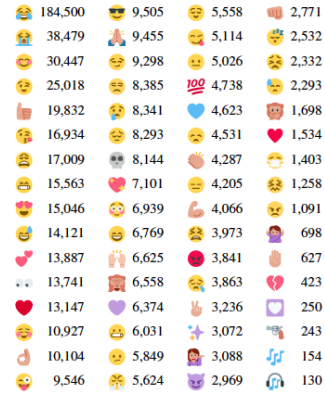


Figure 1: All 64 emoji labels, and number of conversations labeled by each emoji.

The final claim she makes in her article is that empathetic AI will increase transparency and help answer the pertinent question to explain how it arrived in reaching a certain decision. She opined that the empathetic AI, personalized to each customer's needs will help drive greater transparency in its decision making process.

While on some levels, I do agree with these notions, I cannot help but consider the possibility of an empathetic AI also misusing the vulnerability and private information of a human being, upon the tweaks of the corporations involved in harnessing data.

We shall see in the next section, why emotion generation is still not sophisticated enough to make grounding claims.

### 3 Emotion Generation

Zhou et al. in their paper "MOJITALK: Generating Emotional Responses at Scale", aims to experiment with twitter data that is naturally labelled with emojis. They perform the following three tasks - Collect a large data of text corpus with naturally labelled emojis with the assumption that the emojis convey the embedded emotions in a sentence, use variational autoencoders that help them control the emotion generation process and show in their analyses that the proposed models can successfully generate high quality responses based on the assigned emotions.

Overall they aim to classify 64 emoji labels as shown in figure 1. In the data collection process, they stick to the following norms - (1) Emojis with same skin-tone are considered as one (2) Only English tweets without multimedia is crawled (3)

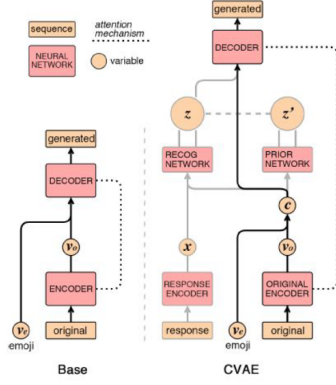


Figure 2: Model architectures

Tweet containing less than 3 words are not included (4) Label responses with emojis (5) Vocabulary size of 20k according to word frequency

They have used attention based sequence to sequence models as baseline with the usage of word embeddings as feature vectors for both sentences and emojis (figure 2). On top of that they used a conditional variational autoencoder (CVAE) for better control of the generated emotions. During training, by  $z$  the recognition network is passed to the decoder and trained to approximate  $z'$  by the prior network. During testing, the target response is absent, and  $z'$  by the prior network is passed to the decoder. For the reinforced CVAE they aim to maximize the expected reward of generated responses.

Model	Perplexity	Emoji Accuracy	
		Top1	Top5
Development			
Base	127.0	34.2%	57.6%
CVAE	37.1	40.7%	75.3%
Reinforced CVAE	38.1	42.2%	76.9%
Test			
Base	130.6	33.9%	58.1%
CVAE	36.9	41.4%	75.1%
Reinforced CVAE	38.3	42.1%	77.3%

Figure 3: Generation perplexity and emoji accuracy of the three models.

In the results, as shown in figure 3, the following takeaways are observed. (1) Higher perplexity score denotes the difficulty level when the model is generating responses (2) CVAE reduces perplexity significantly (3) Mostly unclear explanation of results. For example they did not really define what exactly 'Top1' or 'Top5' was.

The observations regarding the generated sam-

Content	sorry guys , was gunna stream tonight but i 'm still feeling like crap and my voice disappeared . i will make it up to you		
Target Emotion	😞	😞	😞
Base	i 'm sorry you 're going to be missed it	i 'm sorry for your loss	i 'm sorry you 're going to be able to get it
CVAE	hope you are okay hun !	hi jason , i 'll be praying for you	im sorry u better suck u off
Reinforced CVAE	hope you 're feeling it	hope you had a speedy recovery man ! hope you feel better soon , please get well soon	dude i 'm so sorry for that i wanna hear it and i 'm sorry i can 't go to canada with you but i wanna be away from canada
Content	add me in there my bro 🤔		
Target Emotion	😞	👍	❤️
Base	i 'm not sure you 'll be there	i 'm here for you	i 'm not ready for you
CVAE	you know , you need to tell me in your hometown !	you will be fine bro , i 'll be in the gym for you	i can 't wait 😞
Reinforced CVAE	you might have to get me hip hop off .	good luck bro ! this is about to be healthy	i 'm still undecided and i 'm still waiting

Figure 4: Generation samples by the three models.

ples are the following:

(1) It is observed by them that Seq2Seq generates the most frequent utterances in a similar pattern as can be seen in the sections of base in both the contents.(figure 4)

(2) As expected CVAE has more diversity in its generation process. To showcase this disparity, they calculated the type-token ratios of uni-grams/bigrams/trigrams in generated responses as a measure for diversity. The CVAE did better than the base model by a margin of 42.4%.

(3) Reinforced CVAE stack sentences up and its performance is comparable to the CVAE model as seen in figure 4.

Further personal analysis on the paper was done during the class discussion and have been included in the 'Discussion' section later.

## 4 Emojis - Privacy and security

In this report I chose the third source to be an article on emoji privacy by Jack Wallen, primarily because it depicts another face of inducing intelligence in machines, as opposed to the first article I have discussed here.

In his article, Wallen says that although initially emojis evolved to ease means of communication, nowadays emojis encompass emotions, semantic knowledge and sub-textual representations. The fact that companies like Twitter, Facebook and the likes track emoji usage of employees of a company hinders the privacy of the same. These employees can be fall prey to the victim of targeted ads. In other words, his ideas are a direct contradiction to

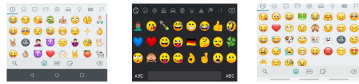


Figure 5: Sample set of recently-used emojis by my friends

the idea of Jun Wu where she says empathetic AI with personalized directions towards a user base can actually felicitate the business models.

Targeting users with advertisements is nothing new in the modern era. There's one example which is cited by Wallen on this account where a Pennsylvanian man was found guilty of threatening his ex-wife by the usage of the ':P' emoticon.

So, in order for AIs to be perfectly empathetic for the benefit of humankind, there has to be an effective combination of scientific prowess and political morality.

## 5 Discussion

For the class discussion session I put two questions forward for the audience:

(1) "Most of us tend to use a specific set of emojis available in the social media we use. - In that case do you think emojis are a good representation of emotions?"

(2) "Do you feel the same emotions are represented differently by emojis, with user privacy ensured? (For example, a private chat application) Also, how can we interpret joint emojis like a 'laughter' and 'thumbs-up' combined?"

On this note I explained that my general intuition is that everyone use a very specific set of emojis to express their emotions. As a pre-work to my presentation I did a small survey on my friends and asked them to send their recent/most-used emoticons. From that I chose a set of 3 emojis which have been included in figure 5.

Its interesting to note that, the set in the middle in figure 5 contains one red heart and one blue heart emoji. Upon inquiring my friend replied he used the blue heart when he expressed his affection on the picture of his wife, who was wearing a blue dress, on their anniversary in social media. In the paper by Zhou et al. the frequency of usage of several emoticons also was surprising because they can have ambiguous contexts as shown in figure 6.

I also put forward in front of the audience, the possibilities of incongruities(4) regarding the cases of sarcastic and ironical tweets. Based on

Possible contexts for this emoji? 13,741

Figure 6: Observation regarding emoji frequencies

this a colleague of mine commented that he generally tends to use a specific set of emojis when he's using social media and that he seldom invests time to find the 'correct' emoji for that particular state of mind of his at that instant. Previously in my presentation, based on the ambiguous representations of emoticons, I commented that emojis can not be used as a singular source of modality to perform emotion analysis, to which he agreed. Another colleague commented that she tends to use the most recently used emojis in her daily life. Although DeepMoji(3) has done well to capture emotion and sarcasm in texts, Zhou et al. leaves the possibilities of incongruency in a tweet.

The main idea of the second question was to put forward the doubt about expressing same emotions differently in two different social medias, one which is in a public domain (like Twitter) and one which apparently has a greater assurance of privacy (like WhatsApp).

The second question did not have a comparable discussion session as the first one due to time constraints. But the overall notion was more inclined to the fact that its probable that people would use different emojis to express the same emotions, if they are assured increased scopes of privacy. So in that aspect, emojis from twitter data does seem to be a tricky choice for emotion generation. As far as usage of joint emojis is concerned, a colleague mentioned when someone texts her something funny she will generally respond with the joint combination of emojis as mentioned in question 2 of my discussion.

## 6 Conclusion

In this report three sources involving incorporation of emotional intelligence in AI, emotion generation and risks of emoji usage have been described with personal insights. A summarized excerpt of the class discussion session post-presentation have also been provided with a background survey. Overall it can be concluded that although there is a long way to go, certain factors combined, we should be able to generate emotions with scraping out the necessary risks involved. But as of now they are too many grey areas involved.



## References

- [1] DAUTENHAHN, K., AND WERRY, I. Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition* 12, 1 (2004), 1–35.
- [2] DIAL, M. *Heartificial Empathy*. Digital Proof Press, 2018.
- [3] FELBO, B., MISLOVE, A., SØGAARD, A., RAHWAN, I., AND LEHMANN, S. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. *arXiv preprint arXiv:1708.00524* (2017).
- [4] JOSHI, A., SHARMA, V., AND BHATTACHARYYA, P. Harnessing context incongruity for sarcasm detection. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (2015), pp. 757–762.
- [5] RAI, A. Explainable ai: from black box to glass box.
- [6] WALLEN, J. The not-so-cute side of emojis: Potential security, privacy, and bandwidth issues.
- [7] WU, J. Empathy in artificial intelligence.
- [8] ZHOU, X., AND WANG, W. Y. MojiTalk: Generating emotional responses at scale. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (Melbourne, Australia, July 2018), Association for Computational Linguistics, pp. 1128–1137.