Assessing Emoji Use in Modern Text Processing Tools

Sunyoung Song

1.
Background

- Prominence of emojis in social media and other instant messaging is growing.
  -> Need to properly handle emojis whenever one deals with textual data is also growing.

Goals

- Our study focuses on tokenization, part-of-speech tagging, and sentiment analysis.
- The results show that many tools have notable deficiencies in coping with modern emoji use in text.

2. Related Work

- **how the broad availability of emojis has affected human communication** (Kaye et al., 2017; McCulloch, 2019)

- **how the use of emoji skin tone modifiers varies on social media** (Robert-son et al. (2020))

- **Many academic studies present new models for particular NLP tasks relating to emojis**.
  - Emoji prediction model for tweets
  - How to extract essential keywords form a tweet using NLP tools
  - Improved part-of-speech tagging based on word clusters
  - Part-of-speech tagger for German social media
  - Understand the use of emojis from a grammatical perspective

3. Experimental Data – (1) Emoji Use in Text

Case 1. Single Emoji

Case 1.1: Single Emoji with Space
Emojis 😀 are a new way of expressing emotions! #emoji

Case 1.2: Single Emoji without Space
Emojis😀are a new way of expressing emotions! #emoji

Case 2. Multiple Emojis

Case 2.1: Multi Emoji Multi Positions
Emojis 😀 are a new way for expressing emotions 😊! #emoji

Case 2.2: Multi Emoji with Space
Another example is having multiple emojis 😀 😀 😀 😀 together in a tweet.

Case 2.3: Multi Emoji Cluster
This gets a little complicated when having multiple emojis 😀😀😀😀 in a tweet without having any spaces in between emojis.

ex. Yahoooo, Funnn

3. Experimental Data – (1) Emoji Use in Text

Case 3. Emojis with Skin Tone Modifiers

- 🏻 Light Skin Tone (e.g. 👍🏻)
- 🏼 Medium-Light Skin Tone (e.g. 👍🏼)
- 🏽 Medium Skin Tone (e.g. 👍🏽)
- 🏾 Medium-Dark Skin Tone (e.g. 👍🏾)
- 🏿 Dark Skin Tone (e.g. 👍🏿)

ex. 🏻 + 👏 -> 👏🏻

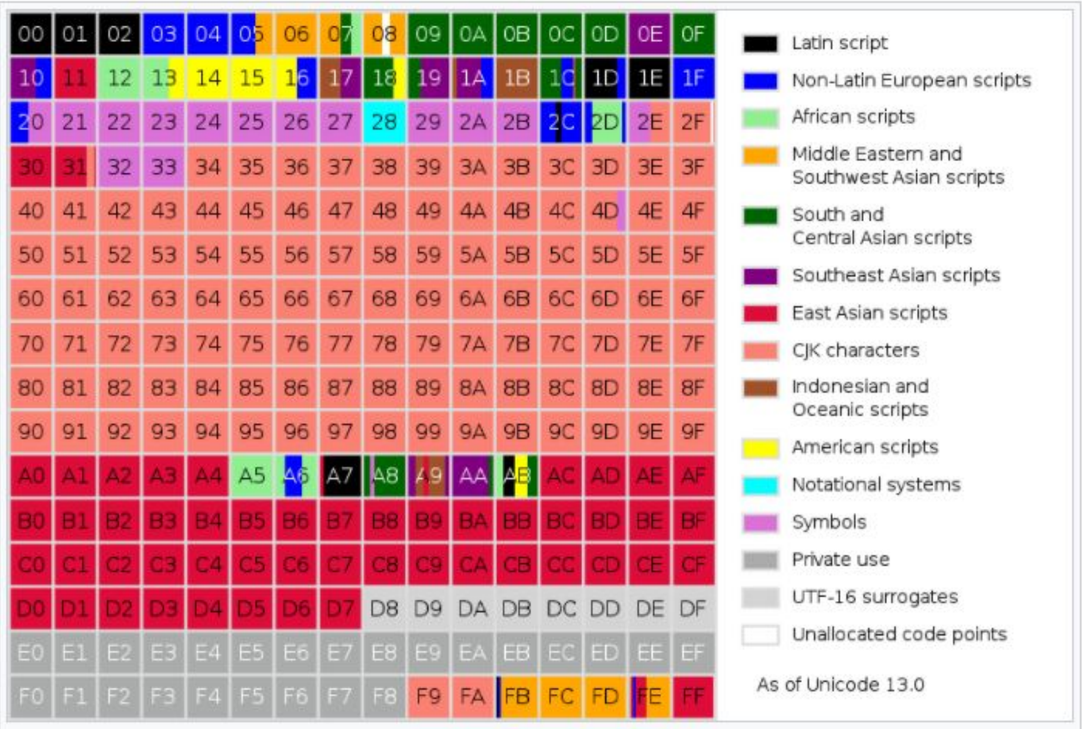
Case 3.1: Skin Tone Emoji with Space
I'm the Face with Tears of Joy emoji 😂. How do you like 👍🏼 me?

Case 3.2: Skin Tone Emoji All Colors
We are all same 👶 👶🏻 👶🏼 👶🏽 👶🏾 👶🏿 but different in skin colors!

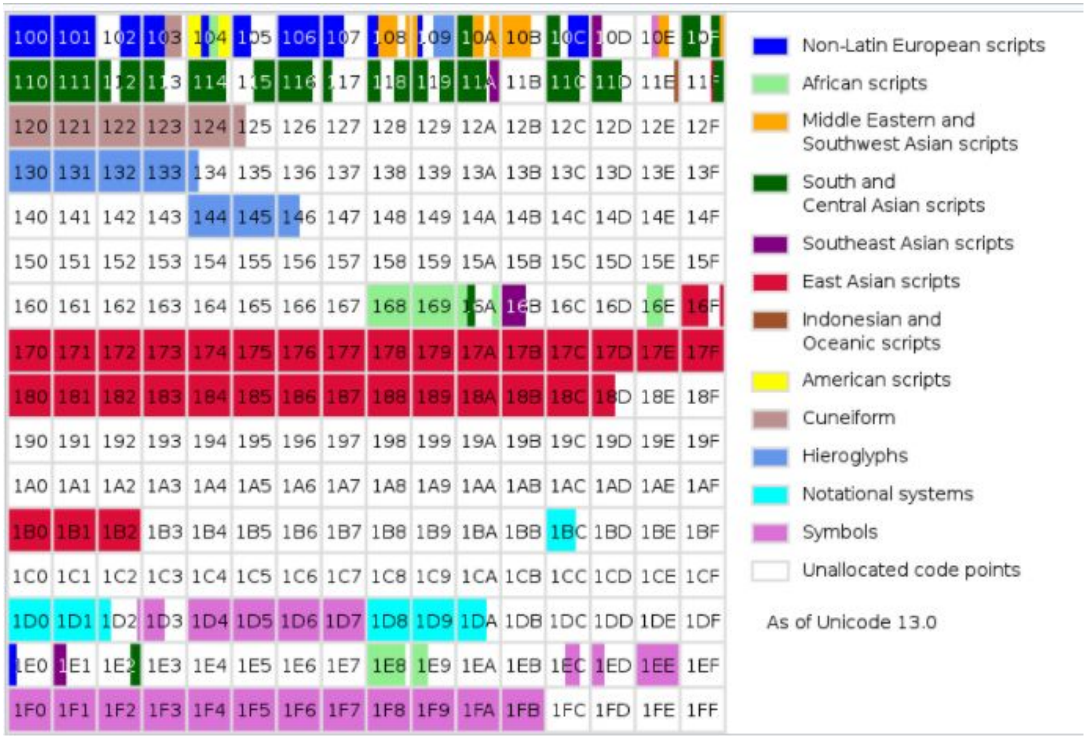
Case 3.3: Skin Tone Emoji Long Sequence
Checking a long sequence of emojis 😂😂😂😂😂😂😂😂 and skin tones 👏👏🏻👏🏼👏🏽👏🏾👏🏿.

3. Experimental Data – (1) Emoji Use in Text

Case 4&5. Emojis from Basic and Supple-mental Planes



BMP



non-BMP

3. Experimental Data – (1) Emoji Use in Text

Case 6. Emojis with Zero Width Joiner (ZWJ)

3. Experimental Data – (2) Tweet Selection

An Example from Twitter
When armed with this everything gets a clean 😂😂😂😂 including the neighbours car 🚗👍🏽😂 ooops #foambath #jetwashing-fun #happydays #everythingclean

- We restricted our search to English language tweets, and made sure that not all tweets simply consisted of URLs or mentions.

- emphasize the motivation for the different cases given in 3-(1)

- Three main tasks were executed to maintain the same tweet selection criteria for all tasks in our experiment.

4. Tokenization

1. Setup

- If multiple emojis occur in a sequence such as "🍰 🥂 🎈 ", they are expected to be treated as individual tokens.

- handling emoji skin tone modifiers adequately

  👋🏻 −> 👋 + 🏻 (x)

- pick a small set of tweets containing ZWJ sequences randomly.

4. Tokenization

2. Results

- We consider 8 libraries for experiments.

Tools	*Task - Tokenization*					
	Case 1 Single Emoji	**Case 2 Multi Emoji**	**Case 3 Skin Tone Emoji**	**Case 4 BMP Plane 0**	**Case 5 non-BMP Other Planes**	**Case 6 Zero Width Joiner**
Gensim	0	0	0	0	0	0
NLTK	80	0	68	40	70	70
NLTK-TT	70	70	0	80	60	0
PyNLPl	50	0	38	30	20	70
SpaCy	100	100	0	100	100	0
SpaCyMoji	100	100	92	100	100	0
Stanza	90	10	68	50	90	40
TextBlob	80	0	68	40	70	70

Table 2: The percentage of success of tools covering all different cases of emojis in tokenization

5. Part-of-speech (POS) Tagging

1. **Setup**

- Excluded Gensim and PyNLPl that do not offer any POS tagging functionality.

- A tool is likely to fail to correctly tag a word or emoji if the emoji is not properly tokenized in the preceding step

  - Experiment based on the output of the integrated tokenizer of the respective tools

    “Emojis😀are”

  - Experiment while considering a unified tokenization as input for all tools

    “Emojis”, “😀”, and “are”.

5. Part-of-speech (POS) Tagging

2. Results

Tools	*Task - Parts-of-Speech (POS) Tagging*						
	Noun 26%	Adjective 22%	Verb ~17.3%	Adverb ~17.3%	Punctuation ~17.3%	Average 100%	Modified Tokenizer
NLTK	100.0	0	0	0	0	26.1	26.1
NLTK-TT	83.3	100	100	0	0	60.9	60.9
SpaCy	66.7	0	100	0	0	34.8	34.8
SpaCyMoji	66.7	0	100	0	0	34.8	34.8
Stanza	83.3	20	100	25	0	47.8	↑ 52.2
TextBlob	83.3	20	100	0	0	43.5	↑ 60.9

Table 3: The percentage of success of tools at labeling emojis with different parts-of-speech. The last column reports the average percentage of success when a modified tokenizer is used.

❓ ‼️ -> marked as nouns

Tools	Tweets	Target Emoji	Expected POS	Default Tokenizer	Modified Tokenizer
Stanza	She kept her 🐕 dog but had to sell her 🐱....	🐕	Noun	🐕 (ADJ) 🐱.... (.)	🐕 (ADJ) 🐱 (NN)
Stanza	I MADE A PICTURE ‼️ ‼️ What do you think ❓ ✨😱	‼️	Punctuation	‼️ ‼️ (.) ❓ ✨ (NNP)	‼️ (NN) ‼️ (.)
TextBlob	I MADE A PICTURE ‼️ ‼️ What do you think ❓ ✨😱	❓	Punctuation	❓ ✨😱 (NNS)	❓ (NNP) ✨ (NNP) 😱 (NN)
TextBlob	Yes, she is😎and I like it	😎	Adjective	is😎and (Verb)	😎 (Adj)
Stanza	**MODIFIED:** She kept her 🐕 but had to sell her 🐱....	🐕	Noun	🐕 (Noun) 🐱.... (.)	🐕 (Noun) 🐱 (Noun)

Table 4: Examples of tweets in which an emoji assumes the role of different parts-of-speech. The last column reports how the tagging accuracy can be improved by utilizing a unified tweet-aware tokenizer across all tools.

6. Sentiment Analysis

1. **Setup**

	Sentiment Predictions		
Sentences	**Only Text**	**Only Emoji**	**Text +Emoji**
They decided to release it 👎🏿	Neutral	–ve	–ve
They decided to release it 😒	Neutral	–ve	–ve
Let's go for it 💃	Neutral	+ve	+ve
My driver license is expired by little over a month 😟	–ve	–ve	–ve
They are going to start a direct flight soon 😢	Neutral	–ve	–ve
They are going to start a direct flight soon 😍	Neutral	+ve	+ve
I'll explain it later 😍	Neutral	+ve	+ve

Table 6: Example sentences with emojis that can moderate the overall sentiment of the sentence

- observe whether the predicted polarity of the original tweet changes in accordance with the polarity of the emojis.

6. Sentiment Analysis

2. Results

Tools	Model	NS	*Emojis* +ve	*Emojis* -ve
NLTK	VADER	100.0	0.0	0.0
TextBlob	PatternAnalyzer	100.0	0.0	0.0
TextBlob	NaiveBayesAnalyzer	100.0	0.0	0.0

Table 5: Accuracy (in percentage) of different tools at predicting sentiment scores of neutral sentences alone (NS) or neutral sentences along with positive (+ve) or negative (-ve) emojis

Examples of tweets in which an emoji assumes the role of different parts-of-speech

7. Discussion

Tools	SE	GE	STE	BMP	ZWJ
AllenNLP	✓	✓	✗	✓	✗
Gensim	✗	✗	✗	✗	✗
NLTK	✓	✗	✓	✓	✓
NLTK-TT	✓	✓	✗	✓	✗
PyNLPl	✓	✗	✓	✓	✓
SpaCy	✓	✓	✗	✓	✗
SpaCyMoji	✓	✓	✓	✓	✗
Stanza	✓	✗	✓	✓	✗
TextBlob	✓	✗	✓	✓	✓

Table 8: An overview of popular text processing NLP tools and their emoji support. Single Emoji (SE), Group Emoji (GE), Skin Tone Emoji (STE), Basic Multilingual Plane (BMP) Plane 0 Emoji, Zero Width Joiner (ZWJ) Emoji

Emoji	Nearest Neighbour Emojis
Clapping Hands (Regular) 👏	👌 👍 👐 👊 🙌
Clapping Hands (Light) 👏🏻	👐🏻 👌🏻 👍🏻 💪🏻 👊🏻
Clapping Hands (Medium Light) 👏🏼	👐🏼 👌🏼 👍🏼 [👏🏽] 👊🏼
Clapping Hands (Medium) 👏🏽	👐🏽 👌🏽 👍🏽 [👏🏼] 💪🏽
Clapping Hands (Medium Dark) 👏🏾	👐🏾 👌🏾 👍🏾 💪🏾 ✊🏾
Clapping Hands (Dark) 👏🏿	👐🏿 👌🏿 👍🏿 💪🏿 👊🏿
ZWJ Family (Man, Woman, Girl, Boy) 👨‍👩‍👧‍👦	👨‍👩‍👧 👨‍👩‍👧‍👧 👩‍👩‍👧‍👧 👨‍👨‍👧‍👦 👩‍👩‍👦

Table 7: Nearest neighbour (NN) emojis for the Clapping Hands and Family emojis. All nearest neighbours follow mostly the same color tone of the respective emojis except some indicated with [ ].

8.
Conclusion

- it is important to endow NLP tools with emoji support

- Our study demonstrates that there is a notable shortcomings in widely used NLP tool by proving that there is no complete tool to handle all cases.

Than
k
You

감사합니
다.