

# Twitterを用いた絵文字の感性と感情ラベル付与

郷間 豊人<sup>†</sup> 李 吉屹<sup>††</sup> 福本 文代<sup>††</sup> 郷 健太郎<sup>††</sup>

<sup>†</sup> 山梨大学工学部コンピュータ理工学科 〒 400-8511 山梨県甲府市武田四丁目 3-11

<sup>††</sup> 山梨大学大学院総合研究部工学域 〒 400-8510 山梨県甲府市武田四丁目 4-37

E-mail: †{t19cs021,jyli,fukumoto,go}@yamanashi.ac.jp

**あらまし** 絵文字予測タスクとは、入力文から、その内容をよく表している絵文字を予測するタスクのことである。このタスクを用いることで、絵文字の持つ情報量からスラングや微妙なニュアンスの違いをとらえたり、テキストの内容を正確に分類したりすることができる。しかし文書に対して適切な絵文字を予測したり、文書や単語に絵文字をラベルとして付与する研究が増えている一方で、絵文字にラベルを付与する研究は少ない。絵文字に感性ラベルや感情ラベルを付与する利点として、絵文字予測タスクによって絵文字を予測できれば、絵文字に付与された感性ラベルや感情ラベルによって更に細かく文書分類が可能に、またそれによりスラングなどの俗語を含んだ分類の難しい文書も感性分析や感情分析が可能となるが挙げられる。しかし、絵文字にラベルを付与するとき、多くの場合主観で人手によって付与される。したがって本研究では、機械的に絵文字に感性と感情ラベルを付与する手法を提案する。

**キーワード** 自然言語処理, 機械学習, 絵文字, 感性ラベル, 感情ラベル

## 1 はじめに

近年、インターネットの普及によって、オンライン上のコミュニケーションが増えた。それにより SNS など、文字以外の表現方法である絵文字がよく使用されるようになった。絵文字は 2010 年に Unicode 6.0 でサポートされ、日本のみならず世界中で Emoji として広く使用されることとなり [1], 研究対象として注目を集めることとなった。以下、Unicode でサポートされた絵文字 (Emoji) の呼称を絵文字と統一する。

絵文字は文字の一部でありながらそれ以上の情報量を有しており、スラングや微妙なニュアンスの違いをとらえたり、テキストの内容を正確に分類したりすることができる。この絵文字の特徴を活かしたタスクに、絵文字予測タスクというものがある。絵文字予測タスクとは、入力文から、その内容をよく表している絵文字を予測するタスクのことである。このタスクを用いることで、絵文字の持つ情報量からスラングや微妙なニュアンスの違いをとらえたり、テキストの内容を正確に分類したりすることができる。主な活用例として、スマートフォンの予測変換機能が挙げられる。

しかし文書に対して適切な絵文字を予測したり、文書や単語に絵文字をラベルとして付与する研究が増えている一方で、絵文字にラベルを付与する研究は少ない。Felbo らの研究ではモデルが絵文字を予測するが、絵文字にラベルが付与されているわけではない。また Ma らのデータセットは、入力文、もしくはアスペクトと絵文字はよく紐付けされているが、感性や感情などは直接紐付けされているわけではない。ここでは、感性とは positive や negative のような極性のこと、感情とは喜びや悲しみのような気持ちのことを指す。絵文字に感性ラベルや感情ラベルを付与する利点として、絵文字予測タスクによって絵文字を予測できれば、絵文字に付与された感性ラベルや感情

ラベルによって更に細かく文書分類が可能に、またそれによりスラングなどの俗語を含んだ分類の難しい文書も感性分析や感情分析が可能となる点が挙げられる。しかし、絵文字にラベルを付与するとき、多くの場合主観で人手によって付与される。Novak ら [2] は、83 人の評価者によって Twitter にて収集した 160 万件のツイートに対し positive, neutral または negative の感性ラベルを付与することで絵文字にラベル付けした。このような人手でラベルを付与する方法は、機械的に付与する方法と比べて非常に時間がかかる、人的コストが高い、個人の感性に左右されるなどの欠点がある。

本研究では、まず SentiWordNet を用いた単語ごとのスコア算出による文書の感性推定を行い文書に感性ラベルを付与した。次に既存の BERT をベースとしたモデルを用いた文書の感情推定を行い、文書に感情ラベルを付与した。次に文書に付与されたラベルの割合から絵文字にラベルを付与した。また、Ma らのデータセットを用いて実際に絵文字に感性と感情ラベルを付与し、付与されたラベルの妥当性をそれぞれ評価した。

## 2 関連研究

Eisner ら [3] は、word2vec [4] と同じ方法で絵文字を空間ベクトルに埋め込む手法を提案した。それにより word2vec を使用するタスクの精度を絵文字の埋め込みによって向上させることに成功した。Felbo ら [5] は、絵文字予測タスクを行うことができるモデル DeepMoji を提案した。このモデルは、16 億の膨大な量の学習データを学習したモデルで、絵文字の仕様における微妙なニュアンスの違いを捉えることができる、また、ファインチューニングすることでほかのタスクにも応用することができ、文脈を理解し感情や皮肉を読み取ったり、似たような文章でも明確に区別したりすることができる。Ma ら [6] は大規模な絵文字予測データセットが存在しないことに注目し、ツイー

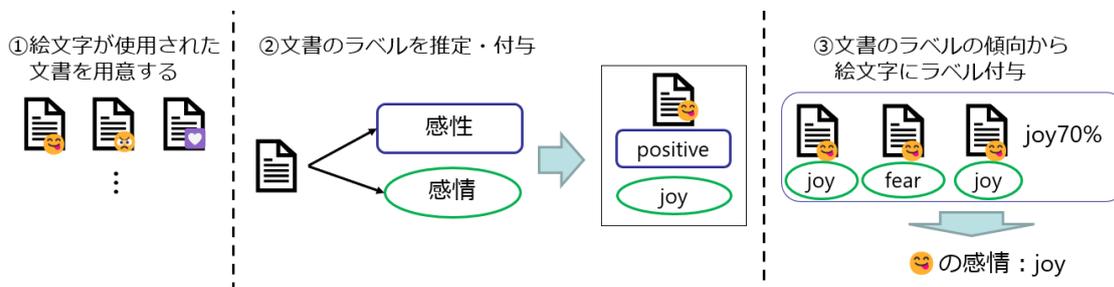


図1 文書へのラベル付与の流れ。

トに対してマルチクラス及びマルチラベル、この手法においてアスペクトと呼ばれている単語または単語列に対してマルチクラスの絵文字をラベルとして付与した大規模な絵文字予測データセットを作成した。このデータセットと、異なるタスクである感情分析や形式分類などのタスクのデータセットとマルチタスク用のモデルでマルチタスクを行うことで、それらのタスクの精度を向上させることができた。Novak ら [2] は、83 人の評価者によって Twitter にて収集した 160 万件のツイートに対し positive, neutral または negative の感性ラベルを付与し、各感性における各絵文字の出現割合を算出した。これにより絵文字に感性スコアを割り当て、絵文字の感情辞書を作成した。また positive なツイートほど絵文字が出現しやすいことを発見した。また Mohammed ら [7] は、人々が言葉だけでは表現しにくいときや、都市名などの固有名詞に情緒的な表現を持たせたい際に絵文字を使用し感情を表現することに着目し、Novak らが作成した絵文字の感情辞書を使用して Twitter の感情分析を行った。その結果、感情分析において絵文字を考慮することで、全体的に positive な文脈の表現力が negative な文脈より相対的に向上することを発見した。

### 3 提案手法

本手法では、絵文字が使用された文書の感性と感情を機械的に予測し感性・感情ラベルを付与し、その統計から絵文字に感性・感情ラベルを付与する。提案手法の流れを図1に示す。

#### 3.1 絵文字を含む文書の準備

絵文字付きの文書は、Ma らが作成したデータセットを使用する。Ma らが作成したデータセットは、文書に付与された絵文字のラベルがマルチクラスのものやマルチラベルのものがあるが、本手法では文書にラベルを付与し、そこからそれぞれの絵文字にラベルを付与するという流れをとるためどちらのデータセットも使用できる。マルチクラスのデータセットにおいて、文書中に異なる絵文字が複数個出現している場合は絵文字の種類だけ文書をコピーし、それぞれの文書に異なる絵文字のラベルを付与している。その特性上、同じ文書に異なるラベルを付与してしまう可能性を考慮し、本研究ではマルチラベルのデータセットを使用する。

データセットに含まれる絵文字の種類は 64 個 [4] となっている。

これは、unicode office の 2021 年絵文字使用頻度<sup>1</sup>より上位 64 個の絵文字を使用している。文書中に同じ絵文字が複数回使用されている場合は、1 つの絵文字につき出現回数は 1 回とみなしている。文書に含まれるハッシュタグや URL、リツイート、メンションは、その内容に関係なく [HASHTAG] や [URL], RT [MENTION]:, [MENTION] のようにトークン化され、ノイズとなるのを防いでいる。またツイートに含まれる URL は添付画像もしくは外部サイトを示している場合が考えられるが、添付画像の場合ツイートの内容を補完している可能性が高いため、URL が含まれるツイートは削除する。最終的な文書数は 8,930 文書である。

#### 3.2 文書の感性ラベル付与

Ma らのデータセットの文書への感性ラベル付与は、SentiWordNet [8] を用いる。SentiWordNet とは、WordNet と呼ばれる意味辞書をもとに作成されたデータセットで、WordNet が提供するデータベース上で動作するセンチメントの尺度を追加機能として持つものである。synset と呼ばれる同義語のグループに、0 から 1 までのスコア *Pos* と *Neg* が割り当てられており、*Pos* が positive 度合い、*Neg* が negative 度合いを示す。

まず Ma らのデータセットの文書に対して形態素解析を行い、単語ごとに分割し品詞を特定する。次にそれぞれの単語を WordNet で検索し該当する synset が存在する場合は SentiWordNet によって品詞と合う synset の *Pos<sub>i</sub>* と *Neg<sub>i</sub>* が付与される。品詞は名詞 (Noun), 動詞 (Verb), 形容詞 (Adjective), 副詞 (Adverb), 前置詞 (Preposition), 接続詞 (Conjunction), 代名詞 (Pronoun), 間投詞 (Interjection) があるが、WordNet は名詞・動詞・形容詞・副詞の 4 つに対応しているため、それ以外の品詞の単語は予め検索を行わない。また、検索を行う際単語が 3 人称や複数などの関係で原型から変化している場合は、原型に直してから検索する。次にこの *i* ごとのスコアを用いて、次の式 (1) によって文書 *s* 全体の感性スコア *Senti(s)* を算出する。*n* は *s* 中の synset が存在する単語の数、*syns<sub>i</sub>* は *i* の属する synset の数を示す。

$$Senti(s) = \sum_{i=1}^n \frac{\sum_{k=1}^{syns_i} Pos_k - \sum_{k=1}^{syns_i} Neg_k}{syns_i} \quad (1)$$

最後に、*s* に対して算出した *Senti(s)* が *Senti(s) < 0* の場

1 : <https://home.unicode.org/emoji/emoji-frequency/>

文書	HAPPY FATHER'S DAY TO ALL THE PAAPAS OUT THERE! I love my Papa Tronco
絵文字	😊
感性ラベル	positive
感情ラベル	joy

表 1 文書に付与されたラベルの例

絵文字	感性 (%)	感情 (%)
😊	positive (86.21)	joy (68.97)
❤️	positive (61.21), negative(33.62)	sadness (41.38)
😱	positive (50.0), negative (44.07)	fear (35.59), surprise (22.88)

表 2 絵文字に付与されたラベルの例.

合 negative,  $Senti(s) = 0$  の場合 neutral,  $Senti(s)$  の場合 positive を付与する.

### 3.3 文書の感情ラベル付与

Ma らのデータセットへの感情ラベル付与は, BERT [9] をベースとした事前学習モデルである emotion-english-distilroberta-base<sup>2</sup>を使用する. BERT とは Attention という仕組みを利用した, 文書を解釈するための自然言語処理における事前学習済みモデルである. emotion-english-distilroberta-base は 6 種類の感情分析データセットによって事前学習されており, ラベルは anger (怒り), disgust (嫌悪), fear (恐怖), joy (喜び), neutral (平静), sadness (悲しみ), surprise (驚き) の 7 種類である. データセットの文書タイプは Twitter, Reddit, 学生のレポート, テレビの対話など様々であり, また感情ラベルそれぞれの文書数は統一されておりそれぞれ 2,811 文書, 合計で約 2 万文書となっている.

本研究では, Ma らのデータセットに合うファインチューニング用のデータセットが存在しないため, emotion-english-distilroberta-base をファインチューニングせずに使用する. 文書をモデルに入力すると, 7 種類のラベルそれぞれの確率が出力される. 本研究では一番確率の高いラベルをそのまま文書に付与した.

予測した文書の感性と感情を, ラベルとして文書に付与する. 最終的なデータセットは, 文書 1 つにつき 1 つまたは複数の絵文字・感性ラベル・感情ラベルが付与されている. 例を表 1 に示す.

### 3.4 絵文字へのラベル付与

64 個すべてにおいて, 1 つの絵文字につき, その絵文字が使用された文書のラベルの統計をまとめ, 全体の割合を 100% として多い割合を占める文書のラベルを絵文字のラベルとして付与する. 例えば, ある絵文字が使用された文書全体について, 感性ラベルは positive が 70%・感情ラベルは joy が 65% であった場合, その絵文字のラベルは positive・joy となる. また, 別の絵文字が使用された文書全体について, 感性ラベルは negative が 90%・感情ラベルは sadness が 35%, fear が 30% であった場合のように, 感情ラベルに複数の傾向が見られるものは, 感性ラベルは negative, 感情ラベルは sadness・fear のように複数ある感情ラベルをすべて挙げ, 付与する. 複数の傾向が見られるとする具体的な判断方法として, まず一番割合の

絵文字	Novak らの結果		本研究	
	positive	negative	positive の割合	negative の割合
😊	0.47	0.25	63.81	32.64
❤️	0.79	0.04	74.55	19.09
❤️	0.69	0.04	77.42	16.59
😍	0.73	0.05	73.81	19.64
😱	0.34	0.44	59.0	34.88

表 3 Novak らとの比較

多いラベルを付与し, 次にそれ以外に割合が 20%以上のラベルがあればそれも付与する. 最後に 15%以上の割合のラベルかつ, さきほど付与されたラベルとの割合の差が 30%未満の場合, そのラベルも付与する.

## 4 実験と考察

### 4.1 実験結果

絵文字にラベルを付与した結果の一部を表 2 に示す.

感性ラベルについては, 絵文字 64 個中すべての絵文字に positive が付与された. この結果から, 絵文字が使用される文書は一般的に positive である傾向にあると考えられる. 文書に付与された感性ラベルについて最も positive の割合が多かった絵文字😊については, 感情ラベルの割合も joy が 8 割を超えるという納得の結果となっているが, 絵文字❤️については, 感情ラベルに sadness が付与されているにも関わらず感性ラベルは positive が 6 割を超える. 一方で絵文字😱のように negative の割合が 4 割を超えた絵文字が 3 つ存在し, また 3 割を超える絵文字もいくつか見られたため, 絵文字によっては negative な文脈でも多く使用されるものもあることがわかる.

### 4.2 評価と考察

#### 4.2.1 Novak らの結果との感性ラベルの評価

感性ラベルの全体的な評価に関しては, positive なツイートほど絵文字が出現しやすいという, 人手でラベルを付与した Novak らの結論と一致したため妥当と考えられる.

また, Novak らが収集したデータの中から頻度順で上位 5 種の絵文字を比較する. 比較したものを表 3 に示す.

スコアの算出方法が異なるため厳密に比較することは難しいが, 絵文字❤️と❤�, 😍に関してはどちらも positive 度合いが非常に高いという点で一致している. また絵文字😍についても positive 度合いが negative 度合いよりも若干高いという点で一致している. しかし絵文字😱に関しては, Novak らの negative 度合いが高いという結果に対して, 本研究では positive 度合いが高いという結果になった. このような差が生まれた要因とし

<sup>2</sup>: Jochen Hartmann, "Emotion English DistilRoBERTa-base". <https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/>, 2022.

データセット	MacroF1			MicroF1		
	最高	最低	平均	最高	最低	平均
(1)	0.5342	0.4680	0.4992	0.6987	0.6438	0.6684
(2)	0.5808	0.5240	0.5447	0.8245	0.7807	0.8029

表 4 評価結果.

て、使用しているツイートデータが異なること、絵文字ごとのツイート数が異なることなどが挙げられるが、一番大きな要因は感性スコアを算出する際すべての synset のスコアを使用してしまったことが考えられる。形態素解析の際に品詞だけでなくどの synset に属する単語かを特定する必要がある。

#### 4.2.2 感情ラベルの機械的な評価

絵文字に付与された感情ラベルを評価するため、感情ラベルを付与した文書を用いて文書分類を行った。データセットとして、(1)8,930 文書に感情ラベルを付与したもの、(2) 絵文字に付与された感情ラベルの傾向と合わないラベルが付与された文書を削除した 5,702 文書の 2 種類を用意した。それぞれ 10 分割したものを検証データとして使用する交差検証を行った。使用するモデルは BERT で、バッチサイズ 24、学習率は 0.00001、ファインチューニングは 5 エポック行った。評価指標には、各クラスの分類性能を平等に評価する指標である MacroF1 と全体の精度を評価する指標である MicroF1 を使用した。結果を表 4 に示す。

評価結果を比較すると、(1) に比べて (2) は MicroF1 が高いことから、絵文字に付与された感情ラベルは比較的正しいものであると考える。一方でいずれも MacroF1 は MicroF1 より低く、また MicroF1 ほど差がない理由として、ラベルごとのデータ数に差がある点が挙げられる。例えば (1) において joy が付与された文書は 4,308 文書だが、anger が付与された文書は 1 割未満の 359 文書しかない。これにより joy のような頻度の高いラベルを付与しやすくなり、anger のような低頻度のラベルを予測することができなくなると考えられる。

#### 4.2.3 感情ラベルのケーススタディ

実験によって絵文字に付与された感情ラベルと異なる感情ラベルが付与された文書に関しては、原因がモデルによる誤った予測か、もしくはその絵文字が特殊な使われ方がされているか考察を行った。その結果、次のような要因が考えられた。

- 傾向と合わないほとんどの文書において、モデルの予測ミスが原因と考えられる。例えば “Emotionally drained after 3 nights of . Looking forward to next season already! 😞” という文書に sadness が付与されていたが、後半の文章を読む限り joy が適切であると考えられる。
- 皮肉表現が用いられている場合が考えられる。例として “Two planned holidays cancelled in the space of 5 days frigg-ing marvellous 🌈” という文書を挙げる。絵文字 🌈 は感情ラベルに joy が付与されているが、この文書は sadness が付与されている。休暇がなくなってしまったことに対する皮肉を絵文字 🌈 で表現していると考えられる。
- 絵文字について言及されている場合が考えられる。例として “When you kinda like someone and then you find out

they have a girlfriend like 😞” という文書を挙げる。この文書は絵文字で文の状況を表している、絵文字なしだと意味がつかめない。このような文書に対しては絵文字なしではラベルを付与することは困難である。

## 5 おわりに

本研究では、絵文字付き Twitter データセットを用いて絵文字に感性と感情のラベルを機械的に付与する手法を提案した。

今後、まずデータセットやモデルを見直し、ラベルの予測を改善する。またそれを改善したのち、文書へ付与されたラベルをもとに統計をとり絵文字にラベルを付与するが、その際絵文字に対する傾向が正しいかを確認するために、協力者に対してアンケートをとる。内容は、傾向に則った絵文字の感性・感情のラベルと、その通りのラベルが付与された任意の文書を見せ、ラベルごとに正しいかどうかを評価するものと、特殊な使用方法であると判断したものに同様に絵文字・ラベル・文書を見せ評価するものである。

## 謝 辞

本研究の一部は、科研費 20K11904 の支援を受けて実施したものである。

## 文 献

- [1] Qiyu Bai, Qi Dan, Zhe Mu, and Maokun Yang. A systematic review of emoji: Current research and future perspectives. *Frontiers in psychology*, 10:2221, 2019.
- [2] Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. Sentiment of emojis. *PloS one*, 10(12):e0144296, 2015.
- [3] Ben Eisner, Tim Rocktäschel, Isabelle Augenstein, Matko Bošnjak, and Sebastian Riedel. emoji2vec: Learning emoji representations from their description. *arXiv preprint arXiv:1609.08359*, 2016.
- [4] Tomáš Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies*, pages 746–751, 2013.
- [5] 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.
- [6] Weicheng Ma, Ruibo Liu, Lili Wang, and Soroush Vosoughi. Multi-resolution annotations for emoji prediction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6684–6694, 2020.
- [7] Mohammed Shiha and Serkan Ayvaz. The effects of emoji in sentiment analysis. *Int. J. Comput. Electr. Eng.(IJCEE.)*, 9(1):360–369, 2017.
- [8] Stefano Baccianella, Andrea Esuli, Fabrizio Sebastiani, et al. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec*, volume 10, pages 2200–2204, 2010.
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.