

テキスト・マルチモーダル埋め込みによる Emoji 意味空間の可視化 Visualizing Emoji Semantic Space via Text and Multimodal Embeddings

奥村 太一¹⁾ 南條 浩輝¹⁾
Taichi Okumura Hiroaki Nanjo

1 はじめに

Emoji は、現代のテキストコミュニケーションにおいて欠かせない表現手段となっている。SNS やメール、ビジネスチャットなどでは、emoji を活用することで文字情報だけでは伝えきれない微妙な感情やニュアンスを補足したり、直截的すぎる表現を和らげたりすることができる。

近年急速に普及した ChatGPT などの生成 AI においても、しばしば emoji を含む出力が得られる。このことは、AI と人間のインタラクションにおいても emoji が人間同士のコミュニケーションと同様の役割を果たしていることを示唆している。

しかし、AI が人間と同じように emoji の意味や感情的ニュアンスを把握できているのかについては、現時点では十分な検証がなされていない。もし人間と AI で emoji の捉え方にずれがあると、AI との対話において意思疎通の齟齬が生じたり、誤解や不要な感情的摩擦を引き起こす可能性がある。したがって、AI が emoji をどのように捉えているかを定量的に分析し、その内部表現を明らかにすることは、人間と AI の円滑なコミュニケーションを実現する上で重要なステップである。

本研究は、こうした emoji の意味的役割や構造を自然言語処理モデルの観点から分析することを目的とするものである。具体的には、テキストおよびマルチモーダル埋め込みを用いてその意味空間を可視化し、emoji が AI モデルにどのように捉えられているかを考察する。

2 関連研究

Emoji のうち、顔の表情を表すもの (facial emoji) について、人間による感情評価を報告した研究がいくつかある。Jaeger et al. (2017)^[1] は、アメリカ人の成人を対象に 33 個の Apple Color Emoji の感情価 (valence) と覚醒度 (arousal) を self-assessment manikin (SAM) によって評価させた。その結果、感情価に比べて覚醒度の分布はばらつきが小さく、互いの間には弱い正の相関関係が見られた。また、感情価の分析から emoji をポジティブ・中立・ネガティブの 3 つのクラスターに分類することができたという。

Kutsuzawa et al. (2022a)^[2], 2022b^[3] は、Twitter (現・X) の提供する Twemoji から 74 個の画像を選んで日本人の成人に提示し、感情価と覚醒度を評定させた。その結果、感情価がニュートラルになるにつれて覚醒度は低くなるという U 字型の分布が見出され、これらの組み合わせから emoji が 6 つのクラスターに分類されたことを報告している。

心理学における感情研究では、コア・アフェクト理論 (Russel et al., 1999)^[4] や感情の輪理論 (Plutchik, 2001)^[5] のように、人間の感情が 2 次元の円環構造で整理で

きるという説が広く支持されてきた。これに対し、Kutsuzawa et al. (2022a)^[2] は、Emoji のような静的な画像では筋肉の緊張や姿勢の変化といった動的な特徴を表現することが難しく、ニュートラルで高覚醒度の領域がうまくカバーされないため円環状に配置されなかった可能性があるとして述べている。

3 方法

3.1 埋め込みモデル

本研究では、emoji の意味的特徴を多角的に捉えるため、表 1 に示す 4 種類の埋め込みモデルを用いて比較分析を行った。

OpenAI の text-embedding-3-large は高次元かつ高性能な多言語対応モデルであり、汎用的な意味表現を捉える能力が高いとされる。MiniLM (paraphrase-multilingual-MiniLM-L12-v2) は軽量ながらも効率的な意味ベクトルを生成できるモデルで、計算資源の制約下でも有用である。LaBSE (Language-agnostic BERT Sentence Embedding) は BERT ベースのアーキテクチャを用いており、特に言語間の意味の一貫性に配慮して設計された多言語対応モデルである。一方、CLIP (japanese-stable-clip-vit-l-16) はテキストと画像の両方を入力可能なマルチモーダルモデルであり、視覚的意味の反映にも強みを持つ。

3.2 Emoji の埋め込み

Kutsuzawa et al. (2022a)^[2] と同じく、顔の表情を表す emoji のうち感情表現が比較的明確であるとされる 74 個を分析の対象とした。テキスト埋め込みモデルには、これらの emoji の Unicode 表現を入力して分散表現を得た。また、CLIP モデルには、Unicode で定義された emoji に対応する Twemoji 画像を 224 × 224 ピクセルの PNG 形式に変換したものをを入力し、分散表現を得た。

3.3 感情語の埋め込み

感情語と emoji の対応関係を把握するために、江川他 (2019)^[6] による 28 個の感情語 (日本語) についても、テキストをそれぞれのモデルに入力して分散表現を得た。

3.4 埋め込みの分析

取得した埋め込み行列について、以下の分析を行った。ただし、MiniLM と LaBSE については複数の emoji で全く同一の埋め込みを返すことがあり、そうしたケースについては分析から除外した。

Emoji の分類 Kutsuzawa et al. (2022a)^[2] が感情価と覚醒度から emoji を分類しているのに合わせ、emoji の埋め込みを UMAP によって一旦 2 次元に削減し、これに階層的クラスター分析 (ユークリッド距離、Ward 法) を適用して emoji を 6 つのクラスターに分類した。この分類結果と Kutsuzawa et al. (2022a)^[2] で報告されている分類結果との一致度を、調整ランド指数 (adjusted Rand index: ARI) と正規化相互情報量 (normalized mutual

1) 滋賀大学データサイエンス学部

