

日本語を母語としない留学生向けの入力内容に応じて 絵文字を推薦する Web アプリケーションの開発

Development of an Emoji Recommendation Web Application for International Students Using Input Text

梅澤 良輔, 佐々木 整

Umezawa RYOSUKE, Sasaki HITOSHI

拓殖大学工学部

Faculty of Engineering, Takushoku University

Email: ryo0127@etil.cs.takushoku-u.ac.jp

あらまし：近年、絵文字が広く利用されているが、日本語非母語話者が文脈に適した絵文字を選択することは容易ではない。著者らは、SNS 上の語句との共起関係に基づいた推薦を行うことで、日本語非母語話者でも自信を持って感情を表現できると考えている。そのための具体的な取り組みとして、SNS 投稿データの分析を行い、入力内容に応じて絵文字を推薦する Web アプリケーションの開発を行った。その結果、推薦候補の中に適切なものが含まれる割合は 80.2% という高い精度を示し、本システムの有効性が確認された一方で、文化やスラングへの対応が今後の課題として明らかとなった。

キーワード：絵文字、絵文字推薦、共起分析、留学生支援

1. はじめに

スマートフォンの普及に伴い、メッセージングアプリを通じたテキストベースのコミュニケーションは我々の生活に不可欠なものとなっている。しかし、対面でのコミュニケーションとは異なり、言語情報のみに依存するコミュニケーションでは、送信者の意図した感情やニュアンスが受信者に正しく伝わらない原因となり、コミュニケーションの齟齬を生む要因となっている。絵文字 (Emoji) は非言語情報を補完する有効な手段として広く普及している。しかし、文献 (1) では日本人学生と外国人留学生の間で絵文字の解釈に差異があることを明らかにしている。特に日本語を母語としない留学生などにとって、適切な絵文字を選択し、意図を正確に伝えることの難しさを指摘している。また、文献 (2) では日本人の大学生を対象とした調査により、一つの絵文字が「幸福」や「悲しみ」といった単一の感情だけでなく、「その他」を含む複数の感情を同時に表す場合があることを明らかにしている。このように、絵文字は文脈によってその解釈が変化しうるため、日本語非母語話者がその場の文脈に即した適切な絵文字を選択することは困難である。そこで本研究では、テキストコミュニケーションにおける円滑な感情伝達を支援するため、入力された日本語文章に応じて最適な絵文字を推薦するシステムの開発を行う。

2. データ収集と品詞抽出

システム開発のため、絵文字と単語の共起確率を計算するためのデータベースを構築した。分析対象となるデータセットとして、SNS プラットフォーム Bluesky から収集した日本語の投稿データを使用した。先行研究に基づき選定された顔文字である 😊(U+1F600 GRINNING FACE) から 😊(U+1F644

FACE WITH ROLLING EYES)までの 69 種類の絵文字を対象とし、各絵文字が含まれる投稿は API を通じて 1,000 件ずつ、合計 69,000 件収集した。収集した投稿データに対し、形態素解析を行った。文脈を構成する要素として「名詞」「動詞」「形容詞」の 3 つの品詞に着目し、それぞれの品詞について抽出を行い、品詞の抽出を行う際には、基本形での収集を行った。抽出された各単語について、特定の絵文字と共に起する確率を算出した。算出された確率は、品詞ごとに、Google スプレッドシートに格納した。

3. 提案システム

分析データを基盤として、ユーザが入力したテキストに対してリアルタイムで絵文字を推薦する Web アプリケーションを開発した。

ユーザが Web インターフェース上のテキストエリアに任意の日本語文章を入力し、検索ボタンを押下すると、入力されたテキストに対して Python の形態素解析ライブラリである Janome を適用し、文を単語単位に分解した上で、「名詞」「動詞」「形容詞」を抽出する。次に、抽出された各単語について、事前に構築されたデータベースを参照し、入力文から抽出されたすべての単語に対し、それぞれの単語がその絵文字と共に起する条件付き確率の総和としてスコアを算出する。そして、算出されたスコアに基づき、69 種類の絵文字を降順にソートする。システムはこのうち上位 5 件を最適な推薦候補として選定し、Web 画面上にボタン形式で提示する(図 1)。最後に、提示された候補の中からユーザが自身の意図に合致する絵文字を選択すると、システムはその選択をデータとして記録する。具体的には、「操作実行時のタイムスタンプ」「入力されたテキスト」「抽出された単語リスト」「システムが提示した候補リスト」、お

より「ユーザが最終的に選択した絵文字」の一連の情報が、Google Sheets API を介してログ保存用の Google スプレッドシートに自動的に追記される。



図 1 Web アプリケーションの実行画面

具体的な実行例として、複数の入力文を用いた場合の動作を述べる。第一に、「猫が可愛くて最高に幸せ」と入力した場合である。この入力に対し、形態素解析によって「猫」「可愛い」「最高」「幸せ」といった単語が抽出される。システムはこれらの共起確率を合算し、😊 (U+1F606) 等を推薦する。第二に、「試験に落ちてしまって悲しい」と入力した場合である。ここでは、抽出された「試験」「落ちる」「悲しい」という単語に基づき、😢 (U+1F622) 等が推薦される。第三に、「手伝ってくれてありがとう」という入力例である。抽出された「手伝う」「ありがとう」という単語に基づき、😊 (U+1F60A) 等が推薦される。これらの実行例から、本システムは多様な文脈において適切な絵文字を推薦し、円滑な感情伝達を支援する機能を有していることが確認できる。特に、言語的なニュアンスの把握が困難な非母語話者にとって、データに基づいた客観的な推薦は、適切な絵文字の選択に有効な手段となる。

4. 有用性の評価

開発したシステムが、日本語非母語話者の感情伝達支援ツールとして有効であるかを検証するため、日本語を母語としない留学生 20 名を対象とした評価実験を行った。実験手順として、まず被験者に本システムを用いて任意の日本語文章を入力させ、システムが推薦した上位 5 件の候補を確認させた。その際、自身の入力意図や感情に合致する絵文字があればそれを、合致するものがなければ画面上の「なし」ボタンを選択するよう指示し、推薦結果が適切か否かを都度二値で判断させるタスクを行った。評価の結果、まずタスクにおいて、推薦した絵文字の中に適切な絵文字が存在した割合は 80.2% であった。これは、本システムが文脈に基づき高い精度で推薦を行えていることを示している。

次に、本システムのユーザビリティを確認するため、実験終了後にシステムは使いやすかったか（操作性）、今後もこのシステムを利用したいか（利用意向）について 5 段階評価（1:全くそう思わない ~ 5:

強くそう思う）で回答を得て、各項目での平均値と標準偏差を求めた（表 1）。また、その評価理由や具体的な使用感について自由記述による回答を求めた。

表 1 各被験者によるシステム評価結果

	操作性	利用意向
平均値	3.90	3.80
標準偏差	1.22	1.29

その結果、「操作性」は平均 3.90 と高い評価を得た。自由記述においては、多くの被験者が画面構成のシンプルさを支持しており、認知的負荷の低い UI 設計が要因といえる。特に、推薦候補数を 5 件に限定した UI 設計は、選択肢過多によるユーザの迷いを防ぎ、迅速な意思決定を促した要因として評価された。この簡潔なインターフェース設計が、二値判定というタスクの遂行を阻害せず、結果として高い操作性評価に寄与したといえる。「利用意向」は平均 3.80 と高水準となった。80.2% という高い割合で適切な推薦が行われていることに加え、高い操作性（3.90）がユーザの入力を支援したことが、高い利用意向に繋がったと考えられる。

5. おわりに

本研究では、ソーシャル・ネットワーキング・サービス（SNS）上の投稿データを収集・分析し、その共起情報を用いて、入力内容に応じた絵文字推薦システムを開発した。日本語を母語としないユーザを対象とした評価実験の結果、推薦した絵文字の中に適切なものが含まれる割合は 80.2% という高い精度を示した。加えて、操作性と利用意向においても高い評価が得られ、Web アプリケーションがテキストコミュニケーションにおける感情伝達を支援するツールとして有効であることが確認された。

今後の課題として、推薦精度のさらなる向上を目指し、学習データの多様化や若者言葉への対応を進める必要がある。また、UI の改良による利便性向上に加え、ログ収集機能によって蓄積されたデータやユーザーからのフィードバックを用いて Google スプレッドシート内に記録されているデータを動的に更新する「適応型システム」の構築を目指す。これにより、文化背景による絵文字の解釈の差異に柔軟に対応し、常に変化するコミュニケーションの形に寄り添う Web アプリケーションへと発展させていきたい。

参考文献

- 佐々木整、斎藤裕美、岡本俊一、倉田芳弥。 “日本人学生と外国人学生による絵文字感情表現の比較調査”, 2024 年電子情報通信学会総合大会講演論文集, D-15-28 (2024)
- Li, C. Sasaki, H. and Saito, H. “What emotions do emojis (emoticons) express: a survey of Japanese university students”, Proceedings of the 8th International Conference on Education and Social Sciences (ICEASS 2024) (2024)