

2022 年度

修 士 論 文

**BERT を用いた
日英間の絵文字の文化差検出**

指導教員：村上 陽平

立命館大学大学院 情報理工学研究科
博士課程前期課程 情報理工学専攻
計算機科学コース

学生証番号：6611210056-4

氏名：松岡 勇樹

BERT を用いた日英間の絵文字の文化差検出

松岡 勇樹

内容梗概

絵文字をテキストに組み込むことで、書き手のテキストに感情やイメージを付与するため、読み手は書き手の意図を容易に理解できる。そのような性質から、チャットコミュニケーションにおける絵文字の使用は、メッセージの書き手と読み手の意思疎通を円滑にする。短い文章で素早くやりとりを行うチャットコミュニケーションと相性が良いため、チャット文中の絵文字の使用が定着している日本語圏や英語圏では日常的に使用されている。

しかしながら、機械翻訳を用いる多言語チャットコミュニケーションにおいて、絵文字の解釈が文化圏で異なるため、絵文字がそのまま挿入されると、翻訳先言語では異なる解釈をされる可能性がある。書き手が原文に込めた感情と同時に使われた絵文字の関連が翻訳によって失われると、読み手は書き手の意図を正しく汲み取れず、コミュニケーション齟齬の原因となる。絵文字を正しく翻訳するために、言語間の絵文字の文化差を検出する必要がある。

そこで、本研究では、絵文字の文化差を検出するために、人が絵文字に抱く感情を定量的に算出し、日本語と英語における絵文字に抱く感情の文化差検出法を提案する。具体的には、絵文字を含むテキストと感情が付与されたテキストを用いて、絵文字に付与された感情を分散表現で表し、その相関から日本語と英語における絵文字に抱く感情の文化差を検出する。本研究では、絵文字に抱く感情を絵文字感情と呼ぶことにする。本手法の実現にあたり、取り組むべき課題は以下の2点である。

絵文字感情モデルの構築

絵文字感情を定量的に算出するためには、絵文字と感情の関連をテキストから抽出する必要がある。人間の表情を示す絵文字は感情と関連付けやすいが、乗り物や動物などの絵文字は感情と関連付けにくい。絵文字の付与されているメッセージの本文を活用して、どのような絵文字に対しても感情の定量化を行うモデルが必要である。

絵文字感情の文化差検出手法

絵文字の文化差の評価を人手で行うと、個々の人間の文化知識が評価に影響されるため、正確な文化差を検出することができない。人手による評価や閾値を用いず、統計的に検出する手法が必要である。

1つ目の課題に対しては、後続絵文字予測モデルと感情予測モデルを構築した。また、各モデルを構築するためのデータセットを生成した。後続絵文字予測モデルを構築するためのデータセットは、**Twitter** から絵文字を含むツイートを収集し、学習のためのデータ整形を行った。感情予測モデルを構築するためのデータセットは、感情タグ付きコーパスから感情値の最も高い文章を抽出した。また、**BERT** の事前学習モデルをファインチューニングし、テキストから後続絵文字と感情によって文章分類を行うタスクを学習させた。

2つ目の課題に対しては、感情タグ付きコーパスを用いて、感情が明確な文章を抽出し、後続絵文字予測モデルに入力した。また、収集したツイートから絵文字を含む文章を抽出し、感情予測モデルに入力した。それぞれの出力結果から、絵文字感情ベクトルを日本語と英語で生成し、スピアマンの順位相関係数で相関を調べた。本研究の貢献は以下の通りである。

絵文字感情モデルの構築

後続絵文字予測モデルは、テキストに後続する絵文字を **1,064** 種類の絵文字から予測でき、感情予測モデルは、テキストに付与された感情を **8** 種類の感情（喜び、悲しみ、期待、驚き、怒り、恐れ、嫌悪、信頼）から予測できた。日本語の後続絵文字予測モデル、英語の後続絵文字予測モデル、日本語の感情予測モデル、英語の感情予測モデルの **accuracy** は、それぞれ **0.15158521**, **0.14533208**, **0.365**, **0.43** であった。

絵文字感情の文化差検出手法

生成した絵文字感情ベクトルをスピアマンの順位相関係数で評価した結果、後続絵文字予測モデルを用いて生成した絵文字感情ベクトルについて、有意水準 **5%** で有意差が認められた絵文字は **497** 種類だった。そのうち、**492** 種類に正の相関があり、**5** 種類に負の相関があった。感情予測モデルを用いて生成した絵文字感情ベクトルについて、有意水準 **5%** で有意差が認められた絵文字は **303** 種類だった。そのうち、**302** 種類に正の相関があり、**1** 種類に負の相関があった。

Cultural difference detection of Emoji between Japanese and English by using BERT

Yuki Matsuoka

Abstract

Readers can easily understand the meaning of text that uses Emoji. The use of Emoji in chat facilitates communication between writers and readers. In Japanese and English, many people often use Emoji.

However, Emoji is not translated in multilingual chat communication. It may be interpreted differently in the target language. If relationship between an original text and Emoji is lost through translation, readers will not be able to understand the intent of the message. It is necessary to detect cultural differences in Emoji in order to translate them correctly.

Therefore, in this study, we propose a method for detecting cultural differences in feelings towards Emoji in Japanese and English. We use the text containing Emoji and the text with emotions to create a vector of emotions attached to Emoji. Based on the degree of similarity, we detect cultural differences in the feelings of Emoji in Japanese and English. There are two issues to be addressed in realizing this method.

Constructing emotion of Emoji prediction models

In order to vectorize Emoji emotions, it is necessary to extract the relationship between Emoji and emotions from the text. Therefore, we need to build the Emoji prediction model and the emotion prediction model to extract the relationship between Emoji and emotions. We need a model that can quantify emotion of any Emoji by using the message with Emoji.

Detecting cultural difference of emotion of Emoji

If people evaluate the cultural difference of Emoji, the cultural knowledge of them is affected by the evaluation. The accurate cultural difference by people cannot be detected. It is necessary to develop the statistical detection method that does not use manual evaluation or thresholds.

For the first task, we construct the Emoji prediction model and the emotion prediction model. In addition, we generate datasets for building each model. The dataset for constructing the Emoji prediction model is composed tweets

containing Emoji from Twitter and process the data for machine learning. The dataset for constructing the emotion prediction model is composed from the corpus tagged best emotions. In addition, we fine-tune the pre-trained model of BERT and train it to classify sentences from texts by subsequent Emoji and emotions.

For the second task, we use the corpus tagged emotion to extract sentences with clear emotions and input them to the Emoji prediction model. In addition, sentences containing Emoji are extracted from tweets and input to the emotion prediction model. We generate emotion of Emoji vectors from output results, and we examine the correlation by using Spearman's rank correlation coefficient. The contributions of this research are as follows.

Constructing emotion of Emoji prediction models

The Emoji prediction model can predict Emoji which follows the text from 1064 Emoji. The emotion prediction model can predict the emotion given to the text from 8 emotions (joy, sadness, anticipation, surprise, anger, fear, disgust, trust). The accuracies of the Japanese Emoji prediction model, the English Emoji prediction model, the Japanese emotion prediction model, and the English emotion prediction model are 0.15158521, 0.14533208, 0.365, and 0.43.

Detecting cultural difference of emotion of Emoji

We evaluate the emotion of Emoji vectors by using Spearman's rank correlation coefficient. There are 497 Emojis with significant difference at the significance level of 5% for emotion of Emoji vectors generated by using the Emoji prediction model. There are 492 positive correlations and 5 negative correlations. There are 303 Emojis with a significant difference at the significance level of 5% for emotion of Emoji vectors generated by using the emotion prediction model. There are 302 positive correlations and 1 negative correlations.

卒業論文タイトル

目次

第 1 章 はじめに	1
第 2 章 絵文字の文化差	3
2.1 絵文字の感情伝達	3
2.2 絵文字解釈の文化差	4
2.3 絵文字使用の文化差	6
第 3 章 学習データ生成	8
3.1 絵文字ラベル付きテキストのデータセット	8
3.1.1 データ収集	8
3.1.2 データ整形	10
3.2 感情ラベル付きテキストのデータセット	12
3.2.1 テキストと感情の関連	12
3.2.2 感情分析データセット	13
3.2.3 学習データセットの作成	14
第 4 章 モデル構築	16
4.1 自然言語処理モデル	16
4.1.1 BERT	16
4.1.2 事前学習モデル	17
4.1.3 ファインチューニング	18
4.2 後続絵文字予測モデル	18
4.2.1 モデル構築	18
4.2.2 モデル評価	19
4.2.3 実行例	21
4.3 感情予測モデル	24
4.3.1 モデル構築	24
4.3.2 モデル評価	24
4.3.3 実行例	26
第 5 章 絵文字感情の分散表現	29

5.1 後続絵文字予測による分散表現.....	29
5.1.1 絵文字感情ベクトルの生成.....	29
5.1.2 出力結果.....	29
5.2 感情予測による分散表現.....	32
5.2.1 絵文字感情ベクトルの生成.....	32
5.2.2 出力結果.....	32
第6章 評価	36
6.1 スピアマンの順位相関係数.....	36
6.2 後続絵文字予測に基づく絵文字感情の相関.....	36
6.3 感情予測に基づく絵文字感情の相関.....	38
6.4 相関の比較.....	39
6.5 考察.....	41
第7章 おわりに	43
謝辞	45
参考文献	46

第1章 はじめに

近年、情報技術が急速に発達し、インターネットは我々の生活に欠かせないものとなっている。インターネットを使ったサービスの一つである SNS は、時間や場所に縛られることなく他者とのコミュニケーションを可能にし、今日も世界中で使用されている。絵文字は、SNS のチャットコミュニケーションにおいて日常的に使用されている。絵文字を文章に組み込むことで、書き手のメッセージに感情やイメージを付与するため、読み手は書き手の意図を容易に理解できる。絵文字の使用は、チャットコミュニケーションにおける意思疎通を円滑にする効果がある。

機械翻訳を用いる多言語チャットコミュニケーションにおいて、絵文字は翻訳されず翻訳文中にそのまま挿入されるため、翻訳先言語では異なる解釈をされる可能性がある。例として、「😭」の日本語と英語の解釈の違いを挙げる。鼻提灯を出しながら眠っている顔を指しており、日本語圏ではそのように解釈されることが多い。しかし、英語圏では鼻提灯を涙と解釈されることが多く、その場合は悲しみを表現する泣き顔と捉えられる。日本語話者が「ぐっすり眠れて気持ちが良い」という喜びを示すために「ぐっすり眠れた😭」と入力すると、翻訳文「I could slept soundly😭」が英語話者に出力される。英語話者が「😭」を泣き顔と解釈した場合、「I could slept soundly」と結び付かず、英語話者は日本語話者のメッセージの意図を理解することが困難である。このように、書き手が原文に込めた感情と同時に使われた絵文字の関連が翻訳によって失われると、読み手は書き手の感情を正しく理解できず、コミュニケーション齟齬の原因となる。したがって、多言語チャットコミュニケーションにおいて絵文字を使用する場合、絵文字の解釈の違いによるコミュニケーション齟齬を防ぐために、テキストだけでなく絵文字も正しく翻訳される必要がある。絵文字翻訳を実現するためには、言語間の絵文字の文化差を抽出する必要がある。

そこで、本研究では、絵文字の文化差を検出するために、絵文字に付与された感情を定量化し、日本語と英語における絵文字に対する感情の文化差検出法を提案する。具体的には、絵文字を含むテキストと感情が付与されたテキストを用いて、絵文字に付与された感情を絵文字感情ベクトルとして算出し、日英間の絵文字感情ベクトルの相関から、日本語と英語における絵文字感情の文化差を検出する。本手法の実現にあたり、取り組むべき課題は以下の 2 点である。

絵文字感情モデルの構築

絵文字感情を定量的に算出するためには、絵文字と感情の関連をテキストから抽出する必要がある。人間の表情を示す絵文字は感情と関連付けやすいが、乗り物や動物などの絵文字は感情と関連付けにくい。絵文字の付与されているメッセージの本文を活用して、どのような絵文字に対しても感情の定量化を行うモデルが必要である。

絵文字感情の文化差検出手法

絵文字の文化差の評価を人手で行うと、個々の人間の文化知識が評価に影響されるため、正確な文化差を検出することができない。人手による評価や閾値を用いず、統計的に検出する手法が必要である。

以下、本論文では、第 2 章で絵文字の感情伝達と従来の絵文字の文化差検出手法について絵文字解釈の文化差、絵文字使用の文化差について述べる。次に、第 3 章で後続絵文字予測モデルと感情予測モデルの構築に使う学習データについて述べる。続いて、第 4 章で後続絵文字予測モデルと感情予測モデルの構築方法について説明し、第 5 章で作成したモデルを用いた絵文字感情ベクトルの生成方法について述べる。第 6 章でスピアマンの順位相関係数の評価結果と日英の絵文字の文化差の考察を行う。最後に今後の展望や課題について述べて結論とする。

第2章 絵文字の文化差

本章では、チャットコミュニケーションにおける絵文字の感情伝達と既存の絵文字の文化差検出法について説明する。

2.1 絵文字の感情伝達

チャットコミュニケーションにおける絵文字の使用は、メッセージに視覚的な情報を与えるため、書き手と読み手の意思疎通を円滑にする。特に書き手が特定の感情を読み手に伝えたい場合にその感情を示す絵文字を使用することで、読み手による書き手の感情理解は促進される。高橋らは、コンピュータを媒介するコミュニケーション（CMC）における書き手と読み手の間の感情伝達における絵文字の効果について、表情を示す絵文字を使用することで正確に感情伝達が行えることを明らかにしている[1]。同時に、感情伝達の精度は感情の種類に依存し、絵文字の感情伝達効果が絵文字のカテゴリによって異なることを明らかにしている。廣瀬らは、絵文字のデザインによって感情伝達の度合いが異なることを明らかにしており、視覚的情報はチャットコミュニケーションにおける意思疎通に重要であることを示している[2]。

絵文字の視覚的情報伝達の可能性を検証する実験も行われている。宗森らは、絵文字のみでチャットを行うためのシステムを開発し、絵文字のみのコミュニケーションの実験を行っている[3]。「日本人学生の仲良しの友達同士」、「日本人学生の普段喋らない者同士」、「日本人学生と留学生（中国人、マレーシア人、ベトナム人）」のペアでそれぞれ実験を行った結果、いずれのペアもごく単純な会話のやりとりであれば、絵文字を組み合わせただけの文章で 70%以上は通じ合えることを明らかにしている。

しかし、多言語コミュニケーションにおける絵文字活用が進められている一方で、絵文字の文化差を原因としたコミュニケーション齟齬が発生する問題がある。具体的には、絵文字の解釈や使用方法が文化によって異なる場合があり、書き手の絵文字の使用や読み手の絵文字解釈が言語文化独特のものであれば、書き手と読み手で正しい情報伝達が行われず、コミュニケーションの崩壊に繋がる。そのような問題を解決するために、絵文字の文化差を検出する研究が行われている。

2.2 絵文字解釈の文化差

CHO らは、日本とアメリカの子供たちが同じ絵文字解釈を持つかを調べるために、120 個の絵文字の意味を問うアンケート調査を行っている[4]。その結果、「ジェスチャー」、「色と性別」、「時間」、「空間」、「馴染みのもの」、「顔/表情」の 6 つの領域で日本とアメリカの子供間で異なる解釈が生じることを明らかにしている。具体例として、「ジェスチャー」と「色と性別」の解釈差について説明する。

まず「ジェスチャー」について、図 1 の左の絵文字は、日本の子供たちは「まる、オッケー、正解、いいよ、あたり」と解釈し、アメリカの子供たちは「**exercise, jump rope, stretch, dance, aerobics**」と解釈した。図 1 の中央の絵文字は、日本の子供たちは「だめ、ぼつ、はずれ、まちがっている、いいえ」と解釈し、アメリカの子供たちは「**angry, mad, frustrated, selfish**」と解釈した。図 1 の右の絵文字は、両国の子供たちの多くが「話す、**taking**」と解釈した一方で、「ありがとう、お願い」と解釈した日本の子供や「**praying**」と解釈したアメリカの子供がいた。

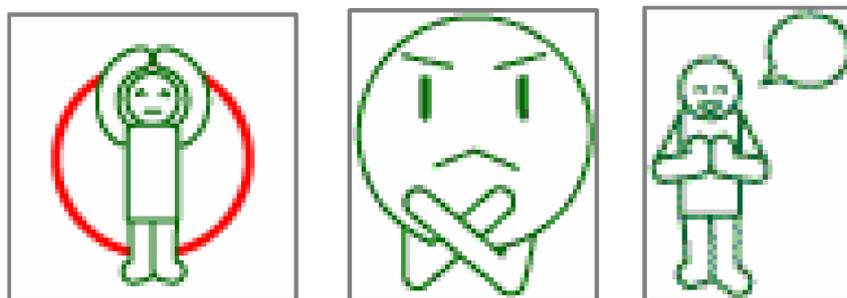


図 1: ジェスチャーの解釈差が顕著に現れた絵文字

出典:[4]

次に、「色と性別」について、図 2 の左の絵文字は、日本の子供たちの 99% が「女性、女の人、女、おかあさん、おねえさん」と解釈したのに対し、アメリカの子供たちの 28% は「**man, dad, boy, male**」と解釈した。図 2 の右の絵文字は、日本の子供たちの 99% が「男性、男の人、男、おとうさん、兄貴」と解釈したのに対し、アメリカの子供たちの 19% は「**woman, mom, big girl, female, old women**」と解釈した。

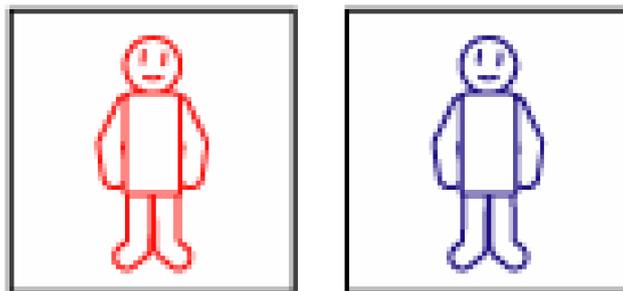


図 2: 「赤は女, 青は男」という色と性別の相関は,
万国共通でない (左は赤, 右は青)

出典: [4]

他の研究では, CHO らは, 絵文字解釈の文化差を検出するために, 日本語と英語の母国語話者に対して, 絵文字の解釈語の違いから文化差評価をさせる実験を行っている[5]. その結果, 絵文字の文化差評価は解釈語の類似性だけでなく, 解釈のばらつきや偏り, 肯定的か否定的か, 固有名詞に関係することを明らかにしている. また, 人手による絵文字解釈の文化差評の問題点として, 評価を行う人間の判断によって揺らぎが発生することを明らかにしている. 例として, 図 3 の絵文字の解釈語について, 日本語は「遊園地」, 英語は「carnival」が最も多かったが, 評価を行った被験者 6 人のうち, 日本人 2 人は異なる解釈であると評価したのに対し, 日本人 1 人とアメリカ人 2 人は同じ解釈であると評価した. 異なる解釈であると評価した被験者は, 「遊園地」と「carnival」は「行事」と「場所」の違いを指摘したのに対して, 同じ解釈であると評価した被験者は, 2 つの単語が連想の関係にあることを指摘している.



図 3: 最も多い解釈語が日本語では「遊園地」,
英語では「carnival」であった絵文字

出典: [5]から一部抜粋

2.3 絵文字使用の文化差

Sharath らは、東洋と西洋の絵文字の使用に関して、両文化における絵文字の使用の基準と違いを定量的に算出する方法を提案している[6]。具体的には、東洋西洋それぞれの絵文字使用の違いを、絵文字の使用頻度、心理言語学的カテゴリに基づく絵文字の使用差、絵文字に含まれる感情と Ekman の基本感情の比較をもとに分析している。

具体的な手法について説明する。Twitter と Weibo から位置情報等を用いて、東洋(アメリカ, イギリス, カナダ)と西洋(日本, 中国)のそれぞれの国で 2014 年に投稿されたメッセージを取得し、テキストデータを形態素解析等でトークン化したコーパスを国ごとに作成する。各コーパスを Word2Vec(モデル: CBoW) で学習し、単語や絵文字を分散表現で表す。また、LIWC によるカテゴリをカテゴリ中の単語の単語ベクトルをもとに平均化したカテゴリベクトルを作成する。絵文字ベクトル, 単語ベクトル, カテゴリベクトルから \cos 類似度を算出し、その値をもとに両文化圏の絵文字の使用の違いを分析する。両文化圏における絵文字の使用頻度については、LIWC カテゴリごとの絵文字使用率ベクトルからスピアマン相関係数を算出する。

LIWC カテゴリごとの絵文字使用率を図 4 に示す。両文化圏の絵文字の使用頻度に関して、スピアマンの順位相関係数は 0.745 であり、両文化間には好まれる絵文字には強い相関がある。また、Unicode Consortium によるカテゴリ(以下 UC カテゴリ)別で絵文字の使用頻度を比較すると、スピアマンの順位相関係数が $0.383\sim 0.807$ であり、使用される絵文字の UC カテゴリにも相関があることを示している。両文化圏とも UC カテゴリが「表情」である絵文字が最も使用されていた。両文化圏の絵文字の使用に関しては、ほとんどの LIWC カテゴリのスピアマン相関係数が正であり、LIWC カテゴリの両文化圏の関連性は高い。その中でも、「Anger」のスピアマンの順位相関係数は 0.738 であり、怒りの表現は両文化圏で類似している。また、絵文字に見られる基本感情に関して、Ekman の基本 6 感情に基づき分析した結果、「怒り」と「喜び」は両文化間で類似性が見られたが、「驚き」、「嫌悪」、「悲しみ」、「恐怖」は類似性が見られなかった。

LIWC		East					West					SCC
Supercategory	Category	1	2	3	4	5	1	2	3	4	5	
Biological processes	Ingest	80%	78%	78%	78%	77%	78%	76%	76%	75%	75%	0.696
	Body	72%	64%	64%	63%	62%	67%	67%	65%	65%	65%	0.473
	Health	72%	72%	69%	69%	61%	63%	62%	62%	62%	61%	0.388
	Sexual	60%	59%	58%	57%	56%	59%	58%	58%	58%	58%	0.133
Cognitive processes	Certain	63%	62%	62%	61%	61%	62%	62%	61%	61%	61%	0.521
	Cause	61%	59%	59%	59%	58%	64%	63%	61%	61%	60%	0.445
	Insight	58%	57%	57%	57%	56%	60%	59%	59%	58%	58%	0.205
	Tentat	57%	57%	57%	57%	56%	58%	57%	57%	57%	57%	0.124
	Discrep	60%	60%	60%	59%	59%	60%	59%	59%	58%	57%	-0.076
Other Grammar	Number	57%	57%	57%	56%	56%	57%	57%	57%	57%	57%	0.161
	Quant	58%	57%	57%	57%	56%	57%	56%	56%	56%	56%	-0.004
Perceptual processes	See	62%	60%	59%	59%	59%	58%	57%	57%	57%	57%	0.492
	Feel	65%	64%	61%	60%	60%	62%	62%	60%	60%	60%	0.431
	Hear	75%	75%	75%	73%	73%	78%	77%	76%	75%	74%	0.414
Personal concerns	Death	63%	63%	62%	61%	61%	64%	62%	62%	62%	61%	0.633
	Home	68%	67%	67%	65%	63%	63%	62%	61%	61%	60%	0.517
	Leisure	67%	65%	64%	63%	63%	62%	61%	61%	61%	61%	0.488
	Money	73%	70%	69%	64%	63%	68%	67%	66%	66%	64%	0.404
	Relig	59%	59%	59%	59%	59%	60%	59%	59%	58%	57%	0.260
	Work	59%	58%	58%	58%	58%	61%	60%	59%	59%	58%	0.139
Psychological Processes	Anger	73%	71%	70%	70%	70%	73%	71%	71%	70%	70%	0.738
	Sad	59%	59%	58%	58%	58%	60%	59%	59%	59%	58%	0.380
	Anx	65%	63%	62%	62%	62%	68%	67%	66%	66%	66%	0.380
	Negemo	61%	60%	60%	59%	59%	66%	66%	65%	63%	63%	0.327
	Posemo	58%	58%	57%	57%	56%	63%	62%	62%	62%	60%	0.220
Relativity	Space	64%	61%	60%	60%	60%	56%	56%	56%	55%	55%	0.360
	Motion	67%	67%	66%	66%	65%	62%	61%	61%	60%	60%	0.350
	Time	60%	59%	58%	58%	58%	59%	57%	57%	56%	56%	0.168
Social processes	Family	70%	69%	68%	66%	64%	68%	67%	64%	64%	64%	0.513
	Friend	61%	59%	58%	58%	58%	58%	58%	58%	58%	58%	0.113

図 4: LIWC カテゴリごとの絵文字使用率

出典: [6]

第3章 学習データ生成

本章では、後続絵文字予測するモデルと感情予測モデルの生成に使用する学習データについて述べる。

3.1 絵文字ラベル付きテキストのデータセット

後続絵文字予測モデルを構築するために、学習データとして絵文字を含むテキストを収集し、テキストと絵文字の一対一対応のデータセットを作成する。また、日本語と英語それぞれの後続絵文字予測モデルを構築するため、日本語と英語のデータセットを作成する。

3.1.1 データ収集

絵文字を含むテキストを収集するためのフィールドとして **Twitter** を使用する。**Twitter** はユーザが全角 140 字以内、半角 280 字以内という比較的短いメッセージを投稿する SNS であり、**DIGITAL 2022: GLOBAL OVERVIEW REPORT** によると 2022 年 1 月時点で 4 億 3600 万人のユーザ数を誇る。**Twitter** に投稿されたツイートを収集するために **TwitterAPI** を使用する。**TwitterAPI** は、検索クエリに条件を指定することでツイートされた日時、言語、投稿者、いいねの数など条件に合ったツイートを取得できる。検索クエリに絵文字を指定することで、絵文字が使用されたツイートを取得できる。

本研究では、Unicode に登録されている 1,146 種類の絵文字を対象に、2022 年 8 月 18 日から 9 月 23 日にかけて各絵文字を含むツイートを収集する。**TwitterAPI** を手軽に使用できるライブラリが複数公開されており、本研究では、python で **TwitterAPI** を呼び出すことができる **tweepy** を使用する。**TwitterAPI** は 1 度の呼び出しで取得できるツイート数に制限があるため、上記の日付について投稿されたツイートの中から、1 時間おきに最大 100 ツイート取得するように設定し、1 日に最大 2,400 ツイート取得できるようにする。取得したツイートからは、投稿された日時、言語、センシティブの内容かどうかの真偽、絵文字を含むテキストを取得する。絵文字「🌀」を含む日本語ツイートから取得する情報の例を表 1 に示す。**TwitterAPI** は絵文字を含むツイートを取得できるが、そのツイート中には絵文字を含むテキストが必ずしもあるとは限らない。また、絵文字単独だけの使用や絵文字が含まれていないテキストも含まれているため、取得したツイートから絵文字を含むテキストのみを抽出する。同時に、完全に重複す

るテキストは排除する。収集時点では、言語を日本語と英語に限定せず、全ての言語のツイートを取得した。以上の条件で収集した結果、**52,470,401** 件の絵文字を含むテキストを取得した。絵文字を含むテキストの取得数の多い上位 10 言語を図 5 に示す。

表 1: 取得したツイート情報の例

daytime	language	sensitive	text
2022-08-20 01:51:25+00:00	ja	False	どれに応募しようかめっちゃ悩む🌀

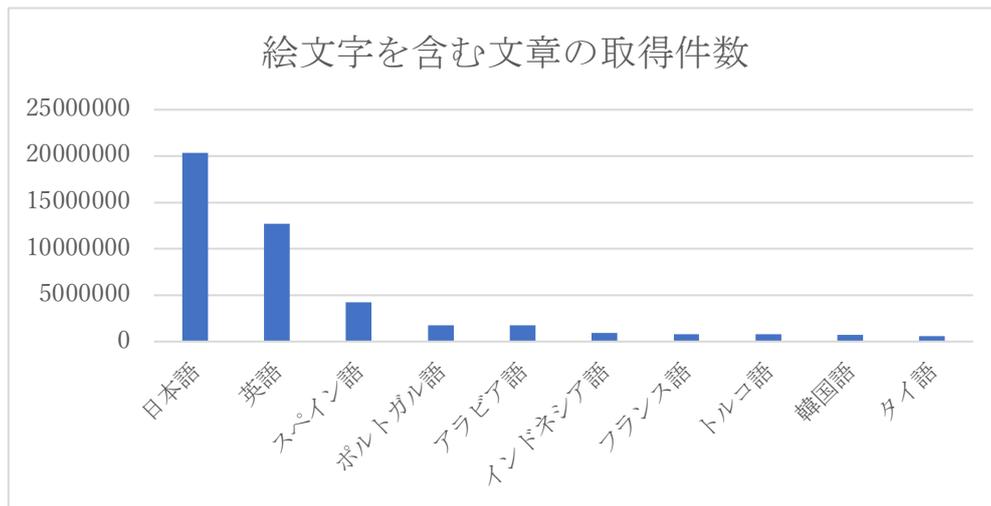


図 5: 言語別の絵文字を含む文章の取得件数

絵文字を含むテキストの取得数が最も多い言語は日本語で**20,345,437**件であった。続いて、英語は**12,678,319**件、スペイン語は**4,242,052**件、ポルトガル語は**1,746,724**件、アラビア語は**1,722,341**件、インドネシア語は**957,787**件、フランス語は**792,642**件、トルコ語は**767,720**件、韓国語は**702,783**件、タイ語は**602,185**件であった。2022年9月の時点における絵文字の使用について、日本語や英語圏のみならずあらゆる言語圏で絵文字が使用されていることがわかる。本研究では、日本語と英語の収集データを用いて、それぞれの絵文字とテキストのデータセットを作成する。データ整形について 3.1.2 で述べる。

3.1.2 データ整形

収集した絵文字を含むテキストを学習データとして扱うために、データの整形を行う。ノイズデータが混入したデータセットを用いて学習すると、低品質なモデル構築の原因となる。したがって、可能な限りノイズデータを除去したデータセットを作成する必要がある。ノイズデータ除去のために行う整形処理の流れを以下に示す。

1. センシティブでないテキストの抽出
2. 一般的な文章の選別
3. 文章中のノイズ除去
4. 文章の文字数制限
5. 絵文字の種類限定

各整形処理について説明する。まず、取得した絵文字を含むテキストからセンシティブでないツイートを抽出する。センシティブな内容に絵文字が使われる場合もあるが、公共のチャットコミュニケーションで使われることは稀であり、多言語コミュニケーション支援の観点では不要である。したがって、本研究では、センシティブな内容と絵文字の関連は考慮しないため、`sensitive` が `True` である絵文字を含むテキストは除去する。

次に、絵文字を含むテキスト中の文字に着目する。取得した文章の言語が日本語であっても、ひらがなを含まないテキストである可能性がある。例として、絵文字「🌀」を含むテキストのうち記号のみのテキストを表 2 に示す。この場合、Twitter 上での言語は日本語であるがテキストは記号のみであるため、学習データとして不適切である。一般的な日本語テキストにはひらがなが含まれているため、ひらがな 50 個のうち 1 個以上含むテキストを抽出した。英語は、一般的な英語テキストには小文字のアルファベットが含まれているため、小文字のアルファベット 26 個のうち 1 個以上含まれるテキストを抽出した。

表 2: ひらがなを含まない日本語文章の例

<code>daytime</code>	<code>language</code>	<code>sensitive</code>	<code>text</code>
2022-08-20 00:59:42+00:00	ja	False	(*≧≦)🌀🌀

続いて、記号を含むテキストに着目する。記号は表 2 のように顔文字を作成に使用する他、表 3 のようにテキストの装飾にも使用される。学習データからモデルを構築する過程でテキストをトークン化するが、テキスト中の記号の種類によっては正常にトークン化されないことがある。表 3 に示す文章中に「〰」の記号が含まれている。記号「〰」は全角一文字であるため、正常にトークン化される場合、未知語、または全角一文字のトークンとして表される。しかし、トークン化するための事前学習モデルによっては、「〰」が 2 つに分割され、正しいトークン数を取得できない。したがって、トークン化によって分割が生じた記号 19 種類をノイズとしてテキスト中から削除した。

表 3: 記号を含む日本語文章の例

daytime	language	sensitive	text
2022-09-07 22:56:25+00:00	ja	False	東北も台風🌀で雨みたい だからきをつけてね •***•.☆

加えて、取得したテキストの文字数制限について説明する。Twitter で投稿できるツイートは文字数制限がかけられている。日本語は全角であるため、日本語でツイートされた文章は基本的に 140 字以内である。しかし、言語情報が正しく入力されていないツイートが混入する可能性がある。例として、言語情報が日本語に設定されているのに対し、アラビア語で記述されたツイートがある。言語情報が正しく入力されていないツイートを自動で完全に除去するには大きなコストがかかるため、テキストの文字数制限を用いる。日本語は 140 字以内、英語は 280 字以内のテキストを抽出することで、言語情報の不正確なテキストを除去する。

最後に、絵文字の限定を行う。使用率の高い絵文字は大量のテキストを取得できるが、使用率の低い絵文字は取得できるテキストが少ない。学習データとして使用する場合は十分な量が必要なため、取得数の少ない絵文字はデータセットから除去する。Unicode に登録されている 1,146 種類の絵文字のうち、日本語と英語どちらも 1000 件以上取得できた絵文字は 1,064 種類である。1,064 種類の各絵文字に対応するテキストを 1,000 件ずつランダムサンプリングし、絵文字とテキストのデータセットとする。絵文字 1,064 種類ごとにテキスト数 1,000

件, 総数 1,064,000 件のデータセットを日英で作成した. 日本語と英語のデータセットの例を, それぞれ表 4, 表 5 に示す.

表 4: 日本語のテキストと絵文字ラベルのデータセットの例

sentence	emoji
本当に遠くても台風の影響受けてる感じですね	
スルメが歯に挟まっていずいとか	
カレーも美味しかったです	
今日のお昼はスタミナドッグで決まりだな	
そうなの！イヴさん！心配してくれてありがとう	

表 5: 英語のテキストと絵文字ラベルのデータセットの例

sentence	emoji
Thank you for everything.	
Had the best Calamari salad	
lunch lunch time	
Marvin literally always got a snack	
what if I don't want them to be just fine	

3.2 感情ラベル付きテキストのデータセット

感情予測モデルを構築するために, 学習データとして感情をタグ付したテキストを収集し, テキストと感情の一対一対応のデータセットを作成する. また, 日本語と英語それぞれの感情予測モデルを構築するため, 後続絵文字予測モデルと同様に日本語と英語のデータセットを作成する.

3.2.1 テキストと感情の関連

テキストと感情の関連には, 書き手がテキストに込めた感情と受け手が読み取る感情の 2 種類がある. 多言語チャットコミュニケーションにおいて絵文字の文化差が生じるのは, 読み手の絵文字から想起する感情が言語によって異なるからである. したがって, 本研究ではテキストに対して受け手が読み取る感情

に着目する。テキストに読み手が読み取る感情を付与するために、人間の感情の指標が必要である。既存の研究では、人間の感情を6種類で示した Ekman の基本6感情[7]（喜び、悲しみ、驚き、怒り、恐れ、嫌悪）や Ekman の基本6感情に（期待、信頼）の感情を加えた、Plutchik の基本8感情[8]がある。テキストと感情の一対一対応のデータセットを作成するために、Plutchik の基本8感情を感情強度としてラベル付した感情分析データセットを用いる。

3.2.2 感情分析データセット

梶原らは、テキストに感情強度をラベル付けした感情分析データセットを公開している[9]。感情分析データセットとは、テキストの書き手の主観的な感情強度のラベルとテキストの読み手の客観的な感情強度のラベルを収集したデータセットである。テキストの書き手と読み手が各8種類の感情を4段階の感情強度でラベル付けを行う。テキストと感情強度ラベルの例を表6に示す。テキスト「タイヤがパンクしてた。。いたずらの可能性が高いんだって。。」に対し、主観は「悲しみ」と「怒り」に3、「驚き」に1の感情強度ラベルを付与している。一方客観は、客観Aが「悲しみ」と「驚き」に3、「恐れ」に2、「怒り」と「嫌悪」に1、客観Bが「悲しみ」と「驚き」に2、客観Cが「悲しみ」と「驚き」に2、「恐れ」と「嫌悪」に1の感情強度ラベルを付与している。この例では、書き手の「怒り」の感情を読み手が検出できていない。収集した主観の感情強度ラベルと客観の感情強度ラベルを比較した結果、書き手の「怒り」と「信頼」の感情を読み手は認識しづらいことを明らかにしている。感情分析データセットを用いて、テキストと感情の一対一対応のデータセットを作成する。

表 6: 感情強度ラベルの例（0：無、1：弱、2：中、3：強）

	喜び	悲しみ	期待	驚き	怒り	恐れ	嫌悪	信頼
主観	0	3	0	1	3	0	0	0
客観A	0	3	0	3	1	2	1	0
客観B	0	2	0	2	0	0	0	0
客観C	0	2	0	2	0	1	1	0

出典: [9]

3.2.3 学習データセットの作成

感情分析データセットからテキストと感情の一対一対応のデータセットを作成する。感情分析データセットは各テキストに 8 種の感情強度ラベルが付与されているため、特定の感情を強く示すテキストを抽出する必要がある。感情分析データセットから日本語と英語の学習用データセットを作成するために行う作業の手順を以下に示す。

1. 客観の感情強度ラベルの総和算出
2. 感情タグ付きテキストの抽出
3. 感情タグ付きテキストの翻訳

各作業について説明する。まず、感情分析データセットからテキストごとに客観の感情強度ラベルの総和を算出する。本研究では、テキストから読み手が読み取る感情を予測するモデルを作成するため、主観の感情強度ラベルは除外する。テキストに客観の特定の感情を付与するためには、客観の 8 種類の感情強度ラベルの総和を求める必要がある。したがって、感情分析データセットのすべてのテキストに対し、客観 A, 客観 B, 客観 C の感情強度ラベルの総和を計算する。

次に、客観の感情強度ラベルの総和からテキストに感情を付与する。テキストの 8 種類の感情強度ラベルの総和から、感情強度ラベルが最も高い感情をテキストの感情として付与する。テキストごとに感情を一意に定めるため、最も高い感情強度ラベルが 1 つのテキストのみを抽出する。最も高い感情強度ラベルが 2 つ以上のテキストは除外する。

最後に、感情タグ付きテキストの翻訳をする。英語の感情タグ付きテキストを作成するために、日本語の感情タグ付きテキストを Google 翻訳によって英語に翻訳する。感情 8 種類ごとにテキスト数 100 件、総数 800 件のデータセットを日英で作成した。日本語と英語のデータセットの例を、それぞれ表 7, 表 8 に示す。

表 7: 日本語のテキストと感情ラベルのデータセットの例

sentence	emotion
参加でめっちゃおもしろい!	joy
4連休にしたけど雨だあ	sadness
横浜方面にでも行こうかな?でも晴れるのか?	anticipation
こんなバファローズ見たことない	surprise
なにしてたんだろうね社会人になるまで	anger
ちょっと地震おおすぎんよ~;;	fear
マジでこの世からコバエ消えて欲しい	disgust
自分の今までの選択に後悔したことはないので	trust

表 8: 英語のテキストと感情ラベルのデータセットの例

sentence	emotion
delicious lemon tea	joy
I want this life to end soon	sadness
let's stay up late	anticipation
I was surprised by the earthquake	surprise
Cats and ladle are somehow	anger
20 million savings or 10 million?	fear
What are you doing? Really stupid.	disgust
We also watch the sports festival.	trust

第4章 モデル構築

本章では、第3章で作成した絵文字とテキストのデータセットと感情とテキストのデータセットを用いたモデル構築について述べる。後続絵文字予測モデルと感情予測モデルをそれぞれ日本語と英語で構築する。

4.1 自然言語処理モデル

テキストから後続絵文字の予測や感情の予測を行うためには、テキストを分散表現で表す必要がある。テキストの分散表現のために使用する自然言語処理モデル BERT の概要、後続絵文字モデルと感情予測モデルを構築するために使用する事前学習モデル、事前学習モデルを絵文字と感情の予測に特化させるためのファインチューニングについて述べる。

4.1.1 BERT

BERT とは 2018 年 10 月に Google の Jacob Devlin らによって発表された自然言語処理モデルである[10]。深層学習モデル Transformer を文頭と文末の双方向に適用することで、文脈を考慮した自然言語処理を可能にしている。従来の自然言語処理モデルには無かった利点として、ラベルのないデータを学習に使えるため、学習コストが大幅に減少したこと、ファインチューニングすることで文章理解や感情分析などの様々なタスクに適用できることが挙げられる。BERT は事前学習とファインチューニングの 2 段階の学習を行い、事前学習は以下の 2 つの手法を組み合わせて行われる。

- Masked Language Model
- Next Sentence Prediction

1 つ目の Masked Language Model は、トークン化したテキストの一部を [Mask] トークンに置き換え、その他のトークンから [Mask] トークンの値を予測するタスクによって学習を行う。例として、テキスト「数学の授業で 1 次関数を習った、数式で直線を表現できる。」の「1 次関数」を [Mask] トークンに置き換えたテキスト「数学の授業で [Mask] を習った、数式で直線を表現できる。」を BERT に入力すると、「数学」、「授業」、「数式」、「直線」という単語から [Mask] トークンの値を予測する。このタスクを繰り返し行うことで、単語の文脈情報を学習する。

2 つ目の Next Sentence Prediction は、入力された 2 つのテキストが連続して

いるかを予測するタスクによって学習を行う。例として、2つのテキスト「[Mask]が無いことに気付いた。週末に[Mask]に行こう。」をBERTに入力すると、連続 IsNext と不連続 NotNext のいずれかで判定を行う。このタスクを繰り返し行うことで、2つのテキストの関係性を学習する。

4.1.2 事前学習モデル

テキストのトークン化と学習には、オープンソースの事前学習モデルを使用する。Hugging Face Hub は自然言語処理に関する事前学習モデルを公開している Web サイトである。Python のライブラリ Transformers を使用することで、Hugging Face Hub に登録されている事前学習モデルをローカルファイルにダウンロードすることなく、手軽に実行することができる。本研究は Hugging Face Hub で公開されている BERT の事前学習モデルを使用する。日本語のモデル構築と英語のモデル構築に使用する BERT の事前学習モデルについてそれぞれ説明する。

日本語の事前学習モデルとして、東北大学の乾研究室が公開している日本語 BERT モデル (cl-tohoku/bert-base-japanese-whole-word-masking) を使用する。モデルは BERT ベースモデルに準拠しており、12 層のレイヤー、768 次元の隠れ状態、12 個のアテンションヘッドで構成されている。学習データとして、2019 年 9 月 1 日時点の日本語版の Wikipedia の記事から抽出したサイズ 2.6, GB, 約 1,700 万件のテキストを用いている。トークン化には、IPA 辞書と MeCab を使用しており、ボキャブラリーサイズは 32,000 である。学習パラメータは BERT ベースモデルと同様に、最大トークン数は 1 インスタンスあたり 512, インスタンスは 1 バッチあたり 256, トレーニングステップは 1,000,000 に設定されている。

英語の事前学習モデルとして、Google が公開している英語 BERT モデル (bert-base-uncased) を使用する。モデルは 12 層のレイヤー、768 次元の隠れ状態、12 個のアテンションヘッドで構成されている。学習データとして、11,038 冊の書籍と英語版の Wikipedia の記事で構成されたデータセットである BookCorpus を用いている。トークン化には、WordPiece を使用しており、ボキャブラリーサイズは 30,000 である。学習パラメータは、最大トークン数は 1 インスタンスあたり 512, インスタンスは 1 バッチあたり 256, トレーニングステップは 1,000,000 に設定されている。

4.1.3 ファインチューニング

テキストから後続絵文字の予測と感情の予測を行うモデルを構築するために、事前学習モデルをファインチューニングする。後続絵文字予測モデルは、テキストから後続絵文字を予測できるようにするために、特徴量をテキスト、ラベルを絵文字とする。テキストに含まれる単語や文脈情報をもとに後続絵文字を予測するタスクを繰り返すことで、テキストと後続絵文字の関係を学習する。感情予測モデルは、テキストから感情を予測できるようにするために、特徴量をテキスト、ラベルを感情とする。テキストに含まれる単語や文脈情報をもとに感情を予測するタスクを繰り返すことで、テキストと感情の関係を学習する。いずれのモデルも第3章で作成したデータセットを使用する。後続絵文字予測モデルには、テキストに絵文字ラベル付けしたデータセットを、感情予測モデルには、テキストに感情をラベル付けしたデータセットを学習データとして使用する。

4.2 後続絵文字予測モデル

後続絵文字予測モデルの概要について述べる。モデル生成方法とモデル性能、実行例をそれぞれ説明する。

4.2.1 モデル構築

テキストと絵文字ラベルの訓練データを用いてモデルを構築する。1,064,000件のテキストを学習するために GPU を使用する。Google が公開している Colaboratory は、無償の GPU 実行環境を制限付きで使用できる。本研究では、Colaboratory Pro を使用した。まず、テキストに絵文字ラベル付けしたデータセットを読み込み、訓練データ、検証データ、テストデータに分割する。データ分割率は、訓練データは 70%、検証データは 15%、テストデータは 15%とし、各データはラベル数が等しい均衡データである。次に、訓練データのテキストをトークン化するために、Hugging Face が提供している Python ライブラリ Transformers を使用する。日本語は BertJapaneseTokenizer、英語は BertTokenizer を用いて BERT 事前学習モデルを読み込む。トークナイザーの BERT 事前学習モデルは日英ともに 4.1.2 で示したモデルを使用する。続いて、訓練データを用いて BERT 事前学習モデルをファインチューニングする。BERT 事前学習モデルを Transformers の BertForSequenceClassification で読み込む。学習評価は、機械学習ライブラリ scikit-learn から sklearn.metrics の accuracy_score, precision_recall_fscore_support を用いて評価関数を定義し、

accuracy, f1 スコア, precision, recall を取得できるようにした. 学習パラメータについて, バッチサイズは 32, エポック数は 1 とし, Transformers の TrainingArguments に定義する. 最後に, BERT 事前学習モデル, 学習パラメータ, 訓練データ, 検証データ, 評価関数を Transformers の Trainer に入力し, train メソッドで学習を実行する.

4.2.2 モデル評価

テストデータを用いて評価した日英の後続絵文字予測モデルの評価結果を示す. 表 9 と表 11 は表 4 のデータセットの例で挙げた 5 種類の絵文字を抜粋している. average は 1,064 種類の絵文字の平均値である. 各スコアについて, 少数第九位以下は切り捨てている. 日本語モデルと英語モデルの評価結果について説明する.

日本語モデルの各スコアを表 10 に示す. accuracy は 0.15158521 であった. 各スコアの平均は, precision は 0.17129028, recall は 0.15158521, f1 スコアは 0.14088384 であった. accuracy が低い原因として, ラベルの数が絵文字の 1064 種類と多いことが考えられる. また, 絵文字の中にはテキストから特徴を得やすいものと得にくいものがあることがわかる. 例として, 「🍛」は, precision, recall, f1 スコアのどれも比較的高い水準を示している. 反対に, 「😞」は, precision, recall, f1 スコアがどれも 0 である. これは, 「🍛」がカレーライスの画像であり, カレーライスに関する文章と共起しやすいことを示している. 「😞」はテキストとの使用に関して汎用性が高く, この絵文字と関連のある単語や文章が発見されなかったことを示している.

英語モデルの各数値を表 12 に示す. accuracy は 0.14533208 であり, 日本語モデルと同程度のスコアを示している. 各スコアの平均は, precision は 0.15704711, recall は 0.14533208, f1 スコアは 0.13327177 である. 日本語モデルと比較すると, 「🍛」と「😞」のスコアは英語モデルの方が高い. また, 「🌀」は precision, recall, f1 スコアがどれも 0 である. これは日本語と異なり, 英語と「🌀」の間に関連のある単語や文章が発見されなかったことを示している.

表 9: 日本語モデルの絵文字 5 種類の評価 (後続絵文字予測モデル)

	precision	recall	f1-score	support
	0.05714285	0.01333333	0.02162162	150.0
	0.06199460	0.15333333	0.08829174	150.0
	0.42792792	0.63333333	0.51075268	150.0
	0.0	0.0	0.0	150.0
	0.00836820	0.01333333	0.01028277	150.0

表 10: 日本語モデルの平均評価と accuracy (後続絵文字予測モデル)

	precision	recall	f1-score	support
average	0.17129028	0.15158521	0.14088384	159600.0
accuracy	0.15158521			

表 11: 英語モデルの絵文字 5 種類の評価 (後続絵文字予測モデル)

	precision	recall	f1-score	support
	0.0	0.0	0.0	150.0
	0.172	0.28666666	0.215	150.0
	0.25454545	0.28	0.26666666	150.0
	0.0	0.0	0.0	150.0
	0.01007556	0.02666666	0.01462522	150.0

表 12: 英語モデルの平均評価と accuracy (後続絵文字予測モデル)

	precision	recall	f1-score	support
average	0.15704711	0.14533208	0.13327177	159600.0
accuracy	0.14533208			

4.2.3 実行例

4.2.1 で構築した後続絵文字予測モデルの実行例を示す。日英のモデルともに入力したテキストから、テキストの後続絵文字の予測結果を 1,064 種類の label と score のセットで出力する。以下の表は、出力結果から score の高い上位 10 種類の絵文字と score を示している。日本語モデルと英語モデルの実行例についてそれぞれ説明する。

まず、日本語モデルの実行例を示す。「雨のせいで憂鬱だ」の出力結果を表 13 に示す。上位 10 件のうち、7 件が表情の絵文字であった。2 番目に高いスコアを示した「🌀」は抽象的な絵文字であるが、日本語では台風や竜巻といった天候や、「ぐるぐる」、「グラグラ」のような擬音、そこから派生し気分や体調が優れないことを示す。また、「憂鬱」に関する絵文字が上位であった。次に、「憂鬱なのは雨のせいだ」の出力結果を表 14 に示す。上位 10 件のうち、4 件が表情の絵文字であった。また、「雨」に関する絵文字が上位であった。2 つの出力結果を比較すると、文末の単語に関連している絵文字が高いスコアを示すことがわかる。「😞」と「🌧️」に着目すると、順位がほとんど逆転している。どちらのテキストも雨が降っていることで気分が憂鬱であることを示し、2 つのテキストの意味的な差異は無い。しかし、2 つのテキストから予測された上位の絵文字の結果が異なることから、予測される絵文字は入力文の文末の単語に関連することがわかる。

次に、英語モデルの実行例を示す。「Because of the rain I'm depressed」の出力結果を表 15 に、「I'm depressed because of the rain」の出力結果を表 16 に示す。最も高いスコアを示した絵文字はいずれも「🌧️」であり、上位 3 件が一致している。2 つのテキストは「because of the rain」と「I'm depressed」の前後を入れ替えたものであり、意味的な差異はない。2 つの出力結果を比較すると、日本語の絵文字予測とは異なり、英語の絵文字予測は単語の順番の影響を受けないことがわかる。一方で「I'm depressed」の出力結果[表 17]は、「depressed」に関連する表情の絵文字が高いスコアを示している。

表 13: 「雨のせいで憂鬱だ」の絵文字予測結果

rank	label	score
1		0.03135275840759277
2		0.024574879556894302
3		0.023456258699297905
4		0.023122793063521385
5		0.021068958565592766
6		0.018649710342288017
7		0.01693812571465969
8		0.015693947672843933
9		0.01553455088287592
10		0.015514026395976543

表 14: 「憂鬱なのは雨のせいだ」の絵文字予測結果

rank	label	score
1		0.04270663484930992
2		0.03346388787031174
3		0.03105887956917286
4		0.0189321581274271
5		0.01868516393005848
6		0.017654182389378548
7		0.017509309574961662
8		0.01670248620212078
9		0.015799112617969513
10		0.015509108081459999

表 15: 「Because of the rain I'm depressed」 の絵文字予測結果

rank	label	score
1		0.24001547694206238
2		0.0357801616191864
3		0.027057237923145294
4		0.022174516692757607
5		0.015762003138661385
6		0.013982408680021763
7		0.011180799454450607
8		0.010398903861641884
9		0.007563451770693064
10		0.007457127794623375

表 16: 「I'm depressed because of the rain」 の絵文字予測結果

rank	label	score
1		0.39384281635284424
2		0.06964964419603348
3		0.044929515570402145
4		0.022497044876217842
5		0.019711708649992943
6		0.014158740639686584
7		0.012974293902516365
8		0.009521981701254845
9		0.00661762710660696
10		0.0065477690659463406

表 17: 「I'm depressed」の後続絵文字予測結果

rank	label	score
1		0.02453548274934292
2		0.020761651918292046
3		0.019944554194808006
4		0.019700944423675537
5		0.019644154235720634
6		0.0194904413074255
7		0.018684400245547295
8		0.01822727359831333
9		0.017848245799541473
10		0.017719389870762825

4.3 感情予測モデル

感情予測モデルの概要について述べる。モデル生成方法とモデル性能、実行例をそれぞれ説明する。

4.3.1 モデル構築

テキストと感情ラベルの訓練データを用いてモデルを構築する。実行環境、データ分割率、トークナイザーは後続絵文字予測モデルの構築と同様のものを用いた。また、モデル構築の手順も後続絵文字予測モデルの構築と同様である。学習パラメータについて、バッチサイズは **32**、エポック数は **20** とした。

4.3.2 モデル評価

テストデータを用いて評価した日英の感情予測モデルの評価結果を表 18 と表 19 に示す。各スコアについて、少数第九位以下は切り捨てている。日本語モデルと英語モデルの評価結果について説明する。

日本語モデルの各スコアを表 18 に示す。accuracy は **0.365** であった。各スコアの平均は、precision は **0.37261911**, recall は **0.365**, f1 スコアは **0.36234854** であった。precision が最も高い感情は **anger** であった。

英語モデルの各スコアを表 19 に示す。accuracy は **0.43** と日本語モデルより

も高くなった. precision が最も高い感情は anger であった. 各スコアの平均は, precision は 0.44732258, recall は 0.43, f1 スコアは 0.43151396 であった.

表 18: 日本語の感情予測モデルの評価結果

	precision	recall	f1-score	support
joy	0.35714286	0.2	0.25641026	25
sadness	0.34375	0.44	0.38596491	25
anticipation	0.42857143	0.48	0.45283019	25
surprise	0.30769231	0.32	0.31372549	25
anger	0.56	0.56	0.56	25
fear	0.25	0.32	0.28070175	25
disgust	0.2962963	0.32	0.30769231	25
trust	0.4375	0.28	0.34146341	25
average	0.37261911	0.365	0.36234854	200
accuracy	0.365			

表 19: 英語の感情予測モデルの評価結果

	precision	recall	f1-score	support
joy	0.46153846	0.48	0.47058824	25
sadness	0.4375	0.28	0.34146341	25
anticipation	0.42857143	0.48	0.45283019	25
surprise	0.41666667	0.4	0.40816327	25
anger	0.6875	0.44	0.53658537	25
fear	0.29032258	0.36	0.32142857	25
disgust	0.375	0.48	0.42105263	25
trust	0.48148148	0.52	0.5	25
average	0.44732258	0.43	0.43151396	200
accuracy	0.43			

4.3.3 実行例

4.3.1 で構築した感情予測モデルの実行例を示す。日英のモデルともに入力したテキストから、テキストの感情の予測結果を 8 種類の感情の label と score のセットで出力する。日本語テキストは日本語モデルに、英語テキストは英語モデルに入力している。4 つのテキスト「雨のせいで憂鬱だ」、「週末公開の映画が楽しみだ」、「I'm depressed because of the rain」、「I'm depressed because of the rain」の出力結果をそれぞれ表 20, 表 21, 表 22, 表 23 に示す。それぞれの表は、8 種類の感情と score を高い順に示している。また、「I'm depressed because of the rain」と「I'm depressed because of the rain」は、それぞれ「雨のせいで憂鬱だ」と「週末公開の映画が楽しみだ」を Google 翻訳で英語に翻訳したものである。4 つのテキストを入力とした日本語モデルと英語モデルの感情予測結果について説明する。

まず、日本語モデルの感情予測結果について述べる。「雨のせいで憂鬱だ」を入力した結果、最も score が高い感情は sadness で、score は 0.9822997450828552 であった。2 番目に score が高い感情は disgust で、score は 0.014789682812988758 であった。また、「週末公開の映画が楽しみだ」を入力した結果、最も score が高い感情は anticipation で、score は 0.6627251505851746 であった。2 番目に score が高い感情は trust で、score は 0.2925938367843628 であった。いずれのテキストも最も score が高い感情の score と他の感情の score と比較すると、最も score が高い感情の score は圧倒的に高い数値を示している。

次に、英語モデルの感情予測結果について述べる。「I'm depressed because of the rain」を入力した結果、最も score が高い感情は sadness で、score は 0.9948464632034302 であった。2 番目に score が高い感情は disgust で、score は 0.002170284977182746 であった。また、「I'm looking forward to the weekend's release of the movie」を入力した結果、最も score が高い感情は anticipation で、0.9965304732322693 であった。2 番目に score が高い感情は trust で、0.0009226319380104542 であった。いずれのテキストも最も score が高い感情の score と他の感情の score と比較すると、日本語モデルと同様に、最も score が高い感情は圧倒的に高い数値を示している。

表 20: 「雨のせいで憂鬱だ」の感情予測結果

rank	label	score
1	sadness	0.9822997450828552
2	disgust	0.014789682812988758
3	trust	0.0011436038184911013
4	fear	0.0005299207987263799
5	joy	0.0004575128259602934
6	anticipation	0.0002884472196456045
7	anger	0.0002556078543420881
8	surprise	0.00023543342831544578

表 21: 「週末公開の映画が楽しみだ」の感情予測結果

rank	label	score
1	anticipation	0.6627251505851746
2	trust	0.2925938367843628
3	joy	0.038659971207380295
4	surprise	0.002456759801134467
5	anger	0.0013891233829781413
6	sadness	0.0010635670041665435
7	fear	0.0005771665601059794
8	disgust	0.0005344259552657604

表 22: 「I'm depressed because of the rain」の感情予測結果

rank	label	score
1	sadness	0.9948464632034302
2	disgust	0.002170284977182746
3	fear	0.0013403274351730943
4	anger	0.0005010499735362828
5	surprise	0.0003867036721203476
6	anticipation	0.000277079816441983
7	joy	0.00026423740200698376
8	trust	0.0002138348063454032

表 23: 「I'm looking forward to the weekend's release of the movie」
の感情予測結果

rank	label	score
1	anticipation	0.9965304732322693
2	trust	0.0009226319380104542
3	anger	0.0006088989903219044
4	fear	0.0005633181426674128
5	joy	0.0005433476762846112
6	surprise	0.0003270218730904162
7	disgust	0.0002743836084846407
8	sadness	0.00022990042634774

第5章 絵文字感情の分散表現

本章では、第4章で生成した後続絵文字予測モデルと感情予測モデルから、絵文字感情ベクトルを生成する方法について述べる。

5.1 後続絵文字予測による分散表現

後続絵文字予測モデルを用いて、1064種類の絵文字の絵文字感情ベクトルを生成する。絵文字感情ベクトルの生成方法と出力結果について述べる。

5.1.1 絵文字感情ベクトルの生成

4.2.1で構築した後続絵文字予測モデルを用いて、絵文字感情ベクトルを生成する。まず、3.2で作成した感情ラベル付きテキストのデータセットを後続絵文字予測モデルに入力する。出力結果から得た1064個のscoreをベクトルに変換する。scoreの高い順に出力されるため、scoreに対応する絵文字のUnicodeコード順に並べ替える。1064次元のベクトルの1番目の要素は「🌀」のscore、2番目は「👁️」のscore、3番目は「👉」のscoreを保存し、これ以降も同様に要素をUnicodeコード順にscoreを保存する。これにより、ベクトルの要素の順番で対応関係が取れる。次に、1064次元のベクトルから感情別に要素の平均を計算する。「joy」の感情ラベル付きテキストの場合、「joy」の感情ラベル付きテキスト100件を後続絵文字予測モデルに入力すると、100個の1064次元のベクトルが得られる。100個のベクトルから要素ごとに平均を感情「joy」と1064種類の絵文字の関係を示す数値とする。そして、ベクトルの各要素を順番に抽出する。

5.1.2 出力結果

5.1.1で生成した日英の絵文字感情ベクトルのうち、30種類の絵文字感情ベクトルを例としてそれぞれエラー! 参照元が見つかりません。と表25に示す。各スコアについて、少数第7位で四捨五入した少数第6位で示している。例として、日本語モデルで生成された「🌀」の絵文字感情ベクトルは、joyは0.00098421, sadnessは0.00251184, anticipationは0.00115605, surpriseは0.00257971, angerは0.00167913, fearは0.00337528, disgustは0.00249058, trustは0.00124383を示している。英語モデルで生成された「🌀」の絵文字感情ベクトルは、joyは0.00117136, sadnessは0.00087146, anticipationは0.00094083, surpriseは0.00112542, angerは0.00078732, fearは0.00140613, disgustは0.00082855, trustは0.00101786を示している。「🌀」の絵文字感情ベクトルは

日英ともに **fear** が最も高かった。

表 24: 日本語の絵文字感情ベクトルの例 (後続絵文字予測モデル)

Emoji	joy	sadness	anticipation	surprise	anger	fear	disgust	trust
	0.000984	0.002512	0.001156	0.002580	0.001679	0.003375	0.002491	0.001244
	0.000543	0.000791	0.000440	0.001773	0.000840	0.002329	0.000805	0.000464
	0.000582	0.001138	0.000525	0.006601	0.000487	0.002115	0.000501	0.000351
	0.000571	0.000491	0.000794	0.000572	0.000289	0.000518	0.000518	0.000369
	0.000411	0.000294	0.000420	0.000497	0.000197	0.000389	0.000251	0.000267
	0.000539	0.000323	0.000590	0.000635	0.000195	0.000382	0.000250	0.000339
	0.000500	0.000384	0.000788	0.000764	0.000232	0.000427	0.000298	0.000318
	0.000567	0.000338	0.000695	0.001000	0.000208	0.000378	0.000230	0.000322
	0.000807	0.000263	0.000543	0.000343	0.000215	0.000247	0.000235	0.000431
	0.000752	0.000450	0.001404	0.000827	0.000261	0.000579	0.000340	0.000499
	0.001248	0.000413	0.001129	0.000934	0.000330	0.000447	0.000366	0.000532
	0.001006	0.000826	0.000904	0.001083	0.000962	0.001104	0.000888	0.000873
	0.000788	0.000468	0.001142	0.000800	0.000298	0.000469	0.000336	0.000654
	0.001022	0.000450	0.001381	0.000771	0.000550	0.000503	0.000440	0.000857
	0.001040	0.000464	0.001373	0.000821	0.000703	0.000553	0.000456	0.000896
	0.000965	0.000350	0.001407	0.000660	0.000450	0.000397	0.000345	0.000696
	0.000902	0.000451	0.000960	0.000631	0.000455	0.000567	0.000444	0.000835
	0.000738	0.000694	0.001183	0.000861	0.000646	0.000679	0.000603	0.000937
	0.000846	0.000642	0.001184	0.000743	0.000430	0.000530	0.000483	0.000712
	0.000647	0.000541	0.000964	0.000690	0.000330	0.000457	0.000411	0.000498
	0.000770	0.000530	0.001026	0.000612	0.000388	0.000416	0.000426	0.000637
	0.000700	0.000562	0.000767	0.000505	0.000302	0.000423	0.000360	0.000506
	0.000655	0.000457	0.000850	0.000543	0.000314	0.000403	0.000355	0.000565
	0.000802	0.000656	0.001191	0.000805	0.000452	0.000552	0.000517	0.000714
	0.000822	0.000713	0.001286	0.000719	0.000470	0.000589	0.000558	0.000763
	0.000660	0.000445	0.000705	0.000409	0.000282	0.000421	0.000398	0.000414
	0.001072	0.002218	0.001342	0.001714	0.002227	0.002213	0.002872	0.001563
	0.000656	0.000508	0.000681	0.000462	0.000305	0.000490	0.000519	0.000404
	0.000789	0.000665	0.000779	0.000555	0.000426	0.000623	0.000684	0.000559
	0.001086	0.001165	0.001131	0.001012	0.000795	0.000892	0.001027	0.000989
	0.000984	0.002512	0.001156	0.002580	0.001679	0.003375	0.002491	0.001244

☀️	0.000543	0.000791	0.000440	0.001773	0.000840	0.002329	0.000805	0.000464
----	----------	----------	----------	----------	----------	----------	----------	----------

表 25: 英語の絵文字感情ベクトルの例 (後続絵文字予測モデル)

Emoji	joy	sadness	anticipation	surprise	anger	fear	disgust	trust
🌀	0.001171	0.000871	0.000941	0.001125	0.000787	0.001406	0.000829	0.001018
🇺🇦	0.000861	0.000842	0.000782	0.000878	0.000469	0.000964	0.000434	0.000781
🖋️	0.001014	0.001075	0.000731	0.003439	0.001325	0.001697	0.001115	0.000591
🏠	0.001147	0.001189	0.000771	0.001073	0.000666	0.001244	0.001213	0.001065
🎨	0.000662	0.000367	0.000453	0.000548	0.000236	0.000352	0.000378	0.000584
🏠	0.000853	0.000540	0.000566	0.000788	0.000287	0.000487	0.000384	0.000625
🏠	0.001014	0.002126	0.000805	0.000993	0.000413	0.000757	0.000492	0.001615
🏠	0.000942	0.001040	0.000792	0.000960	0.000380	0.000734	0.000456	0.000802
🌈	0.001105	0.000839	0.000694	0.000680	0.000558	0.000692	0.000651	0.001291
🏠	0.000977	0.001209	0.000703	0.000782	0.000627	0.001154	0.000876	0.001186
🏠	0.000943	0.000520	0.000955	0.000806	0.000533	0.000607	0.000500	0.000881
🏠	0.001018	0.000549	0.000802	0.001295	0.000671	0.001327	0.000450	0.000782
🏠	0.001206	0.000831	0.000968	0.000754	0.000725	0.000952	0.000658	0.001111
🌐	0.000488	0.000314	0.000661	0.000534	0.000263	0.000634	0.000372	0.001760
🌐	0.000678	0.000433	0.000878	0.000684	0.000355	0.000714	0.000485	0.001906
🌐	0.000541	0.000274	0.000751	0.000476	0.000240	0.000505	0.000279	0.001256
🌐	0.000493	0.000385	0.000665	0.000605	0.000351	0.000528	0.000370	0.000819
🌑	0.001045	0.000876	0.001494	0.000797	0.001121	0.001108	0.000682	0.001054
🌑	0.001081	0.000668	0.001364	0.000740	0.000960	0.000925	0.000542	0.000949
🌑	0.001022	0.000713	0.001123	0.000837	0.000781	0.000966	0.000552	0.000960
🌑	0.000982	0.000606	0.001102	0.000675	0.000794	0.000839	0.000473	0.000790
🌑	0.000696	0.000518	0.001356	0.000807	0.000773	0.000854	0.000396	0.000695
🌑	0.000687	0.000539	0.001145	0.000583	0.000811	0.000772	0.000441	0.000811
🌑	0.001025	0.000751	0.001275	0.000881	0.000966	0.001091	0.000616	0.001005
🌑	0.001275	0.000896	0.001341	0.000872	0.001075	0.001137	0.000793	0.001151
🌙	0.000968	0.000701	0.000822	0.000584	0.000696	0.000866	0.000777	0.000801
🌑	0.001118	0.001889	0.001073	0.001333	0.001868	0.001685	0.002061	0.001101
🌙	0.001159	0.000840	0.000875	0.000690	0.000686	0.000996	0.000866	0.001029
🌙	0.001238	0.000830	0.000892	0.000655	0.000682	0.000969	0.000814	0.000966
😊	0.001363	0.001794	0.001047	0.001200	0.001589	0.001605	0.001630	0.001266

😊	0.001171	0.000871	0.000941	0.001125	0.000787	0.001406	0.000829	0.001018
🌟	0.000861	0.000842	0.000782	0.000878	0.000469	0.000964	0.000434	0.000781

5.2 感情予測による分散表現

感情予測モデルを用いて、1064 種類の絵文字の絵文字感情ベクトルを生成する。絵文字感情ベクトルの生成方法と出力結果について述べる。

5.2.1 絵文字感情ベクトルの生成

4.3.1 で構築した感情予測モデルを用いて、絵文字感情ベクトルを生成する。まず、3.1 で作成した絵文字ラベル付きテキストのデータセットを感情予測モデルに入力する。出力結果から得た 8 個の score をベクトルに変換する。score は高い順に出力されるため、score に対応する感情を joy, sadness, anticipation, surprise, anger, fear, disgust, trust の順に並べ替える。5.1.1 と同様に、ベクトルの要素の順番で対応関係が取れるようにする。次に、8 次元のベクトルから絵文字別に要素の平均を計算する。「😊」の絵文字ラベル付きテキストの場合、「😊」の絵文字ラベル付きテキスト 1000 件を感情予測モデルに入力すると、1000 個の 8 次元のベクトルが得られる。得られた 1000 個のベクトルから要素ごとに平均を計算する。その平均値を絵文字「😊」と 8 種類の感情の関係を示す数値とする。そして、ベクトルの各要素を順番に抽出し、絵文字感情ベクトルを生成する。

5.2.2 出力結果

5.1.1 で生成した日英の絵文字感情ベクトルのうち、30 種類の絵文字感情ベクトルを例としてそれぞれ表 26 と表 27 に示す。各スコアについて、少数第 7 位で四捨五入した少数第 6 位で示している。例として、日本語モデルで生成された「😊」の絵文字感情ベクトルは、joy は 0.13960524, sadness は 0.24968772, anticipation は 0.1450295, surprise は 0.07378176, anger は 0.04354329, fear は 0.32382661, disgust は 0.02233952, trust は 0.00218635 を示している。英語モデルで生成された「😊」の絵文字感情ベクトルは、joy は 0.16917438, sadness は 0.05511942, anticipation は 0.20922682, surprise は 0.13779299, anger は 0.08215152, fear は 0.16306206, disgust は 0.09856338, trust は 0.08490943 で示している。日本語の「😊」の絵文字感情ベクトルは fear が最も高かったが、英語の「😊」の絵文字感情ベクトルは anticipation が最も高かった。

表 26: 日本語の絵文字感情ベクトルの例 (感情予測モデル)

Emoji	joy	sadness	anticipation	surprise	anger	fear	disgust	trust
	0.139605	0.249688	0.145030	0.073782	0.043543	0.323827	0.022340	0.002186
	0.287807	0.231008	0.100268	0.120749	0.016413	0.180498	0.052800	0.010456
	0.192530	0.161949	0.177797	0.121309	0.047776	0.210766	0.039527	0.048346
	0.411662	0.113648	0.156553	0.071943	0.063089	0.069007	0.010763	0.103335
	0.567342	0.090627	0.153957	0.058123	0.006286	0.056594	0.021772	0.045299
	0.539146	0.063294	0.141719	0.106427	0.017087	0.045499	0.029307	0.057521
	0.436167	0.065226	0.092138	0.086418	0.058329	0.062863	0.015691	0.183169
	0.431716	0.116851	0.073680	0.114803	0.045509	0.110852	0.040171	0.066417
	0.364416	0.060956	0.257680	0.051204	0.042132	0.073014	0.041901	0.108697
	0.321470	0.142895	0.187822	0.094400	0.069596	0.108552	0.012735	0.062529
	0.417279	0.075708	0.188491	0.083335	0.049039	0.083626	0.032298	0.070224
	0.292043	0.087527	0.168213	0.136831	0.053907	0.118511	0.100202	0.042767
	0.437192	0.072702	0.234460	0.035121	0.035596	0.070877	0.022534	0.091518
	0.289497	0.063819	0.312605	0.128094	0.034680	0.065786	0.020624	0.084895
	0.291716	0.021387	0.270578	0.171475	0.050311	0.066002	0.073040	0.055491
	0.308827	0.037418	0.328731	0.109086	0.048815	0.055965	0.037637	0.073521
	0.250626	0.074293	0.340277	0.125498	0.075098	0.059566	0.014179	0.060462
	0.320872	0.192416	0.220262	0.056719	0.070394	0.048495	0.016755	0.074086
	0.360812	0.068056	0.197602	0.043292	0.086591	0.081382	0.080054	0.082212
	0.447187	0.135943	0.180014	0.068496	0.027836	0.075064	0.020492	0.044968
	0.389124	0.104101	0.194639	0.122379	0.043769	0.079861	0.048713	0.017415
	0.376137	0.097662	0.258389	0.127684	0.032705	0.024715	0.003520	0.079188
	0.468992	0.089718	0.168150	0.103414	0.016822	0.058496	0.023991	0.070417
	0.328787	0.108634	0.236251	0.077959	0.030274	0.101529	0.028212	0.088353
	0.337893	0.089818	0.157447	0.132353	0.068443	0.103558	0.039613	0.070875
	0.530017	0.101466	0.120631	0.085120	0.029044	0.055251	0.027147	0.051325
	0.182195	0.101621	0.107535	0.062826	0.126289	0.140400	0.180842	0.098293
	0.458276	0.089898	0.178612	0.043172	0.019329	0.073095	0.033312	0.104306
	0.486081	0.091218	0.153824	0.024496	0.044912	0.098619	0.038592	0.062259

😊	0.335232	0.115070	0.154663	0.122280	0.050764	0.073763	0.083352	0.064875
😬	0.139605	0.249688	0.145030	0.073782	0.043543	0.323827	0.022340	0.002186
🌟	0.287807	0.231008	0.100268	0.120749	0.016413	0.180498	0.052800	0.010456

表 27: 英語の絵文字感情ベクトルの例 (感情予測モデル)

Emoji	joy	sadness	anticipation	surprise	anger	fear	disgust	trust
🌀	0.169174	0.055119	0.209227	0.137793	0.082152	0.163062	0.098563	0.084909
🖼️	0.286248	0.044698	0.183605	0.177621	0.073232	0.074899	0.064552	0.095147
🖋️	0.178504	0.031907	0.338481	0.114196	0.144337	0.068360	0.065909	0.058305
🏠	0.253945	0.053778	0.230374	0.103794	0.148701	0.041786	0.043229	0.124394
🏡	0.451365	0.028933	0.154956	0.106694	0.106149	0.022466	0.036452	0.092985
🇪🇺	0.376545	0.028627	0.224146	0.171933	0.061626	0.002105	0.004710	0.130308
🏠	0.324398	0.057399	0.251122	0.053828	0.068915	0.022369	0.075341	0.146628
🏠	0.269615	0.042248	0.232171	0.136322	0.097055	0.067393	0.052791	0.102405
🌈	0.292202	0.029095	0.184127	0.081511	0.100470	0.011499	0.100250	0.200845
🏠	0.211465	0.047765	0.243488	0.126182	0.096684	0.082378	0.034424	0.157614
🌊	0.308481	0.036778	0.239585	0.159679	0.041310	0.028995	0.054667	0.130505
🏠	0.168089	0.029532	0.312074	0.207643	0.083624	0.101436	0.045171	0.052431
🏠	0.294607	0.059327	0.173910	0.094847	0.126377	0.069566	0.044613	0.136754
🌐	0.234078	0.027151	0.226421	0.133237	0.054977	0.044331	0.066302	0.213504
🌐	0.305683	0.045219	0.226973	0.110498	0.050331	0.036202	0.064736	0.160357
🌐	0.182851	0.055020	0.284857	0.155824	0.094192	0.027542	0.054153	0.145560
🌐	0.125158	0.044381	0.427936	0.142964	0.030759	0.070842	0.053720	0.104239
🌑	0.144376	0.044692	0.200666	0.213203	0.071554	0.136314	0.079965	0.109230
🌑	0.211493	0.070905	0.267165	0.198350	0.034801	0.052267	0.055098	0.109921
🌑	0.273852	0.042626	0.250395	0.168344	0.080799	0.031844	0.070225	0.081914
🌑	0.215739	0.031158	0.229748	0.168635	0.077205	0.053976	0.024424	0.199116
🌑	0.205508	0.034398	0.302161	0.170770	0.058185	0.114984	0.043581	0.070413
🌑	0.190944	0.014924	0.274510	0.150665	0.085331	0.065930	0.085407	0.132288
🌑	0.146340	0.087771	0.259080	0.149835	0.084400	0.086115	0.054416	0.132043
🌑	0.199524	0.125435	0.159822	0.196958	0.054486	0.070066	0.048691	0.145019
🌑	0.327290	0.062703	0.224988	0.119405	0.079779	0.034046	0.066129	0.085660
🌑	0.189800	0.106233	0.155711	0.146211	0.130523	0.057559	0.147232	0.066730
🌑	0.281995	0.040208	0.215150	0.128062	0.123131	0.041393	0.039535	0.130525

🌙	0.292416	0.070778	0.213864	0.163054	0.013800	0.058721	0.083765	0.103602
😊	0.165474	0.049296	0.140712	0.203071	0.148468	0.068434	0.115642	0.108903
☀️	0.169174	0.055119	0.209227	0.137793	0.082152	0.163062	0.098563	0.084909
🌟	0.286248	0.044698	0.183605	0.177621	0.073232	0.074899	0.064552	0.095147

第6章 評価

本章では、第5章で算出した各絵文字の感情ベクトルから日英間の文化差を検出する手法について述べる。

6.1 スピアマンの順位相関係数

絵文字感情ベクトルから日英の文化差を検出するために、評価手法としてスピアマンの順位相関係数を用いる。スピアマンの順位相関係数とは、2変数間の順位データから相関を測る指標である。以下の式によって算出できる。

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

d は対応するデータの順位の差、 n は順位データのペアの数で定義される。絵文字感情ベクトルにスピアマンの順位相関係数を適応するために、絵文字感情ベクトルの各要素に順位をつける。各要素を数値の高い順に1から8の数値に置換する。例として、表24の「🌀」に対して各要素に順位をつけると、joyは8、sadnessは3、anticipationは7、surpriseは2、angerは5、fearは1、disgustは4、trustは6となる。

6.2 後続絵文字予測に基づく絵文字感情の相関

後続絵文字予測モデルから生成した絵文字感情ベクトルのスピアマンの順位相関係数を求めた結果、有意水準5%で有意差が認められた絵文字は497種類だった。そのうち、正の相関があった絵文字は492種類、負の相関があった絵文字は5種類であった。正の相関があった絵文字の上位10種類と負の相関があった絵文字5種類をそれぞれ表28、表29に示す。

まず、正の相関を示した絵文字について述べる。正の相関を示した絵文字の上位10件のうち、スピアマンの順位相関係数が最も高い絵文字は「👊」、「😬」、「👊」、「😬」、「👊」であり、スコアはいずれも1.0であった。また、「👊」、「👊」、「👊」、「👊」、「👊」のスコアはいずれも0.97619であった。それぞれのスコアの最も高い感情と2番目に高い感情はいずれの絵文字も共通しており、「👊」はdisgustとanger、「😬」はdisgustとanger、「👊」はdisgustとanger、「😬」はdisgustとanger、「👊」はjoyとanticipation、「👊」はdisgustとanger、「👊」はjoyとanticipation、「👊」はjoyとtrust、「😬」はdisgustとanger、「😬」は

disgust と sadness であった. 上位 10 種類の絵文字のうち, 7種類の絵文字「👎」, 「😞」, 「🐼」, 「😡」, 「👊」, 「😜」, 「😐」が日英ともに disgust と anger の感情を示していた.

次に, 負の相関を示した絵文字について述べる. 負の相関を示した絵文字 5 種類のうち, スピアマンの順位相関係数が最も低い絵文字は「🌸」で, スコアは-0.880952 であった. 続いて, 「💧」は-0.857143, 「/」は-0.761905, 「👤」は-0.738095, 「💥」は-0.714286 であった. それぞれの絵文字についてスコアの最も高い感情と 2 番目に高い感情は, 日本語における「🌸」は anger と disgust, 英語における「🌸」は sadness と anticipation, 日本語における「💧」は trust と joy, 英語における「💧」は disgust と anger, 日本語における「/」は anticipation と trust, 英語における「/」は disgust と anger, 日本語における「👤」は joy と anticipation, 英語における「👤」は disgust と anger, 日本語における「💥」は anger と disgust, 英語における「💥」は trust と joy であった. 5 種類の絵文字の結果から, joy, anticipation, trust のポジティブな感情と anger, disgust のネガティブな感情が日英間では反対に現れた.

表 28: 正の相関があった絵文字の上位 10 種類 (後続絵文字予測)

rank	Emoji	correlation
1	👎	1.00000
2	😞	1.00000
3	🐼	1.00000
4	😡	1.00000
5	👊	1.00000
6	👊	0.97619
7	🎡	0.97619
8	🍡	0.97619
9	😜	0.97619
10	😐	0.97619

表 29: 負の相関があった絵文字 5 種類 (後続絵文字予測)

rank	label	correlation
1		-0.880952
2		-0.857143
3		-0.761905
4		-0.738095
5		-0.714286

6.3 感情予測に基づく絵文字感情の相関

感情予測モデルから生成した絵文字感情ベクトルのスピアマンの順位相関係数を求めた結果、有意水準 5% で有意差が認められた絵文字は 303 種類だった。そのうち、正の相関があった絵文字は 302 種類、負の相関があった絵文字は 1 種類であった。正の相関があった絵文字の上位 10 種類と負の相関があった絵文字をそれぞれ表 30, 表 31 に示す。

まず、正の相関を示した絵文字について述べる。正の相関を示した絵文字の上位 10 種類のうち、スピアマンの順位相関係数が最も高い絵文字は「」, 「」, 「」であり、スコアはいずれも 1.0 であった。また、「」, 「」, 「」, 「」, 「」, 「」, 「」のスコアはいずれも 0.97619 であった。それぞれのスコアの最も高い感情と 2 番目に高い感情について、「」, 「」, 「」, 「」, 「」, 「」, 「」は共通しており、「」と「」は異なった。具体的には、「」は joy と anticipation, 「」は joy と anticipation, 「」は joy と anticipation, 日本語における「」は joy と anticipation, 英語における「」は anticipation と joy, 「」は joy と anticipation, 「」は joy と anticipation, 日本語における「」は joy と anticipation, 英語における「」は anticipation と joy, 「」は joy と anticipation, 「」は joy と anticipation, 「」は anticipation と joy であった。上位 10 種類の絵文字のうち、すべての絵文字が日英ともに joy と anticipation の感情を示していた。

次に、負の相関を示した絵文字について述べる。負の相関を示した絵文字は「」の 1 種類で、スピアマンの順位相関係数は -0.714286 であった。「」のスコアの最も高い感情と 2 番目に高い感情は、日本語における「」は

anticipation と sadness, 英語における「😡」は anger と surprise であった。

表 30: 正の相関があった絵文字の上位 10 種類 (感情予測)

rank	label	correlation
1		1.00000
2		1.00000
3		1.00000
4		0.97619
5		0.97619
6		0.97619
7		0.97619
8		0.97619
9		0.97619
10		0.97619

表 31: 負の相関があった絵文字 (感情予測)

rank	label	correlation
1		-0.714286

6.4 相関の比較

後続予測に基づく絵文字感情の相関と感情予測に基づく絵文字感情の相関を日本語と英語の同一言語同士で比較した結果、有意水準 5% で有意差が認められた日本語における絵文字は 247 種類、英語における絵文字は 237 種類であった。そのうち、正の相関を示した日本語における絵文字は 236 種類、負の相関を示した日本語における絵文字は 11 種類、正の相関があった英語における絵文字は 231 種類、負の相関があった英語における絵文字は 6 種類であった。負の相関を示した日本語における絵文字 11 種類と、英語における絵文字 6 種類をそれぞれ表 32, 表 33 に示す。

まず、負の相関を示した日本語における絵文字について述べる。負の相関を示

した日本語における絵文字のうち、スピアマンの順位相関係数が最も低い絵文字は「😡」で、スコアは**-0.833333**であった。続いて、「🔑」、「🤢」、「🦷」、「🏠」はいずれも**-0.809524**、「👤」は**-0.785714**、「📺」、「😞」はいずれも**-0.761905**、「🔪」、「😜」、「👉」はいずれも**-0.714286**であった。それぞれの絵文字についてスコアの最も高い感情と 2 番目に高い感情は、後続絵文字予測に基づく「😡」は **anger** と **disgust**, 感情予測に基づく「😡」は **joy** と **surprise**, 後続絵文字予測に基づく「🔑」は **anger** と **trust**, 感情予測に基づく「🔑」は **joy** と **anticipation**, 後続絵文字予測に基づく「🤢」は **disgust** と **anger**, 感情予測に基づく「🤢」は **anticipation** と **joy**, 後続絵文字予測に基づく「🦷」は **anger** と **surprise**, 感情予測に基づく「🦷」は **anticipation** と **joy**, 後続絵文字予測に基づく「🏠」は **trust** と **sadness**, 感情予測に基づく「🏠」は **joy** と **anticipation**, 後続絵文字予測に基づく「👤」は **disgust** と **trust**, 感情予測に基づく「👤」は **joy** と **anticipation**, 後続絵文字予測に基づく「📺」は **sadness** と **anger**, 感情予測に基づく「📺」は **joy** と **anticipation**, 後続絵文字予測に基づく「😞」は **trust** と **disgust**, 感情予測に基づく「😞」は **joy** と **anticipation**, 後続絵文字予測に基づく「🔪」は **anger** と **fear**, 感情予測に基づく「🔪」は **joy** と **anticipation**, 後続絵文字予測に基づく「😜」は **disgust** と **sadness**, 感情予測に基づく「😜」は **joy** と **anticipation** であった。

次に、負の相関を示した英語における絵文字について述べる。負の相関を示した英語における絵文字のうち、スピアマンの順位相関係数が最も低い絵文字は「🔪」で、スコアは**-0.809524**であった。続いて、「😞」は**-0.785714**、「😞」は**-0.761905**、「🏠」、「👉」はいずれも**-0.738095**、「zZ」は**-0.714286**であった。それぞれの絵文字についてスコアの最も高い感情と 2 番目に高い感情は、後続絵文字予測に基づく「🔪」は **fear** と **sadness**, 感情予測に基づく「🔪」は **anticipation** と **joy**, 後続絵文字予測に基づく「😞」は **sadness** と **fear**, 感情予測に基づく「😞」は **anticipation** と **joy**, 後続絵文字予測に基づく「😞」は **sadness** と **disgust**, 感情予測に基づく「😞」は **joy** と **anticipation**, 後続絵文字予測に基づく「🏠」は **fear** と **disgust**, 感情予測に基づく「🏠」は **joy** と **anticipation**, 後続絵文字予測に基づく「👉」は **disgust** と **anger**, 感情予測に基づく「👉」は **surprise** と **anticipation**, 後続絵文字予測に基づく「zZ」は **disgust** と **fear**, 感情予測に基づく「zZ」は **anticipation** と **joy** であった。

表 32: 負の相関を示した日本語における絵文字

rank	label	correlation
1		-0.833333
2		-0.809524
3		-0.809524
4		-0.809524
5		-0.809524
6		-0.785714
7		-0.761905
8		-0.761905
9		-0.714286
10		-0.714286
11		-0.714286

表 33: 負の相関を示した英語における絵文字

rank	label	correlation
1		-0.809524
2		-0.785714
3		-0.761905
4		-0.738095
5		-0.738095
6		-0.714286

6.5 考察

日英で正の相関があった絵文字に着目すると、後続絵文字予測に基づくスピアマンの順位相関係数の上位 10 種類の絵文字と感情予測に基づくスピアマンの順位相関係数の上位 10 種類の絵文字に共通する絵文字は無い。また、6.4 で示したように同一言語同士でも正の相関と負の相関を示す絵文字に差がある。こ

の原因として、後続絵文字予測モデルを用いて生成した絵文字感情ベクトルと感情予測モデルを用いて生成した絵文字感情ベクトルは、絵文字感情の判断基準が異なることが考えられる。テキストと絵文字の関係を考えると、テキストの文末に選出される絵文字はテキスト中の単語に関連する場合が多い。一方で、テキストと感情の関係を考えると、テキストに込められた感情はテキスト中の単語に関連しない場合がある。例として、「今日は苦手な上司が病気で休みだった」に対して、単語単位で読むと「苦手」や「病気」という単語から、テキストにはネガティブな感情が込められていると考えられる。しかし、テキストの意味を考えるとポジティブな感情が込められているとわかる。自然言語処理モデルによる感情の予測と人間による感情の予測の間には、単語の前後関係だけでなく背景や状況の理解があるかという点で異なる。したがって、人間によって予測された感情を利用する場合と機械学習によって感情を予測する場合には、絵文字感情の算出方法が異なる。以上の理由から、後続絵文字予測に基づくスピーアマンの順位相関係数の上位 10 種類の絵文字と感情予測に基づくスピーアマンの順位相関係数の上位 10 種類の絵文字が異なると考えられる。

日英で負の相関があった絵文字に着目すると、6 種類の絵文字に絵文字感情の文化差があると考えられる。蜘蛛を表す絵文字「🕷️」について、日本語テキストでは蜘蛛に対する嫌悪感を示し、英語テキストではアメリカンコミックのスーパーヒーローであるスパイダーマンに対する感情を示していると考えられる。ツバ付きの帽子を表す絵文字「🧢」について、日本語テキストでは遊びに出かけることに対する喜びを示し、英語テキストではヒップホップの楽曲「No Cap」から広まった嘘を意味するスラングに関する感情を示していると考えられる。白杖を表す絵文字「🦯」について、日本語テキストでは視覚障害者に対する思いやりの感情を示し、英語テキストでは目が見えない困難さに対する感情を示していると考えられる。手を頭上で合わせている様子を表す絵文字「🙌」について、日本語テキストでは同意に対する感情を示し、英語テキストでは捕まっている様子に対する感情を示していると考えられる。衝突を表す絵文字「💥」について、日本語テキストでは気持ちや思いを発散することに対する感情を示し、英語テキストではおめでたいことに対する感情を示していると考えられる。

第7章 おわりに

本研究では、日英の絵文字の文化差を検出するために BERT を用いた絵文字の文化差検出手法を提案した。まず、モデル構築のための学習データとして、テキストに絵文字ラベルを付与したデータセットとテキストに感情ラベルを付与したデータセットを作成した。作成したデータセットを用いて、テキストから後続絵文字を予測する後続絵文字予測モデルと、テキストから感情を予測する感情予測モデルを構築した。構築した 2 つのモデルを用いて、絵文字感情を定量化した絵文字感情ベクトルを生成した。生成した絵文字感情ベクトルからスピアマンの順位相関係数を算出し、日英の絵文字感情ベクトルの相関を測った。本研究の貢献は以下の通りである。

絵文字感情モデルの構築

後続絵文字予測モデルは、テキストに後続する絵文字を 1,064 種類の絵文字から予測でき、感情予測モデルは、テキストに付与された感情を 8 種類の感情（喜び、悲しみ、期待、驚き、怒り、恐れ、嫌悪、信頼）から予測できた。日本語の後続絵文字予測モデル、英語の後続絵文字予測モデル、日本語の感情予測モデル、英語の感情予測モデルの accuracy は、それぞれ 0.15158521, 0.14533208, 0.365, 0.43 であった。

絵文字感情の文化差検出手法

生成した絵文字感情ベクトルをスピアマンの順位相関係数で評価した結果、後続絵文字予測モデルを用いて生成した絵文字感情ベクトルについて、有意水準 5% で有意差が認められた絵文字は 497 種類だった。そのうち、492 種類に正の相関があり、5 種類に負の相関があった。感情予測モデルを用いて生成した絵文字感情ベクトルについて、有意水準 5% で有意差が認められた絵文字は 303 種類だった。そのうち、302 種類に正の相関があり、1 種類に負の相関があった。

今後の展望として、モデル精度の向上が挙げられる。本研究では、後続絵文字予測モデルの構築に 1,064 種類の絵文字を含むテキストを 1,000 件ずつの計 1,064,000 件の学習データを用いたが、絵文字 1 種類あたりのテキスト数やエポック数を増加させることで予測精度を向上させることができる。感情予測モデルは、感情 1 種類あたりのテキスト数は 100 件であり、計 800 件の学習データを使用したため、感情 1 種類あたりのテキスト数を増加させると予測精度を向上

させることができる。また、学習データに感情タグ付きコーパスを用いたが、読み手がテキストから読み取る感情は個々の人間によって異なるため、可能な限り多くの客観による評価を用いることが望ましい。

加えて、英語の感情ラベル付けデータセットを作成するために、日本語の感情ラベル付きのデータセットの翻訳を行ったため、読み手の母国語が日本語である場合の感情強度が適用されている。同じテキストでも、言語によって受け手に対する感情伝達が異なる場合があるので、読み手を英語が母国語話者にした感情分析データセットを作成すべきである。

最後に、絵文字感情ベクトルの相関から言語による絵文字感情の文化差を検出できるため、絵文字翻訳によって感情伝達を行うコミュニケーションシステムへの応用が期待される。また、絵文字翻訳を実現するために、より多くの絵文字を対象とした文化差検出を行う必要がある。

謝辞

本研究を行うにあたり，熱心なご指導，ご助言を賜りました指導教官の村上陽平准教授に深謝申し上げます。また，普段からお世話になっている社会知能研究室の皆さまに心より感申し上げます。

参考文献

- [1] 高橋直己, 上野舞夕, 浜田百合, 庄司裕子: 絵文字を用いた文章における感情伝達効果に関する研究, 日本感性工学会論文誌, Vol. 21, No. 1, pp. 135-142 (2022).
- [2] 廣瀬信之, 牛島悠介, 森周司: 携帯電話メールによる感情の伝達に顔文字と絵文字が及ぼす影響, 感情心理学研究, Vol. 22, No. 1, pp. 20-27 (2014).
- [3] 宗森純, 大野純佳, 吉野孝: 絵文字チャットによるコミュニケーションの提案と評価, 情報処理学会論文誌, Vol. 47, No. 7, pp. 2071-2080 (2006).
- [4] Cho Heeryon, 稲葉利江子, 石田亨, 高崎俊之, 森由美子: 絵文字コミュニケーションにおけるセマンティクス, 情報処理学会研究報告知能と複雑系(ICS), Vol. 2006, No. 110, pp. 1-8 (2006).
- [5] Cho Heeryon, 石田亨, 山下直美, 稲葉利江子, 高崎俊之, 神田智子: 絵文字解釈における人間の文化差判定, ヒューマンインターフェース学会論文誌, Vol. 10, No. 4, pp. 427-434 (2008).
- [6] Sharath Chandra Guntuku, Mingyang Li, Louis Tay, and Lyle H Ungar: Studying Cultural Differences in Emoji Usage across the East and the West, In Proceedings of the International AAAI Conference on Web and Social Media, Vol. 13, pp. 226-235 (2019).
- [7] Paul Ekman: An Argument for Basic Emotions, Cognition and Emotion, Vol. 6, No. 3-4, pp. 169-200 (1992).
- [8] Robert Plutchik: A General Psychoevolutionary Theory of Emotion, Theories of Emotion, Vol. 1, pp. 3-31 (1980).
- [9] 梶原智之, Chenhui Chu, 武村紀子, 中島悠太, 長原一: 主観感情と客観感情の強度推定のための日本語データセット, 言語処理学会第27回年次大会, pp. 523-527 (2021).
- [10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).