

現代日本語短歌の感情識別への取り組み

黒川 真琳 (指導教員：小林 一郎)

1 はじめに

人の情動は、未だ十分に解明されていない複雑な心理現象である。これまで、情動に関する多くの研究が行われており、脳神経科学の観点からその特性を調査する試みもなされている [1]。一般に、情動は感情と密接に結びついていると考えられ、その要因を明らかにすることは重要である。本研究では、言語芸術の一つである短歌を対象とし、短歌が誘発する情動と特定の感情との関連性を調査する。これにより、抽象的な概念である情動を、具体的な感情の要素を通じて表現することを試みる。

2 研究概要

現代短歌を対象として、短歌が誘発する情動と特定の感情との関連性を調査する。研究概要を図 1 に示す。

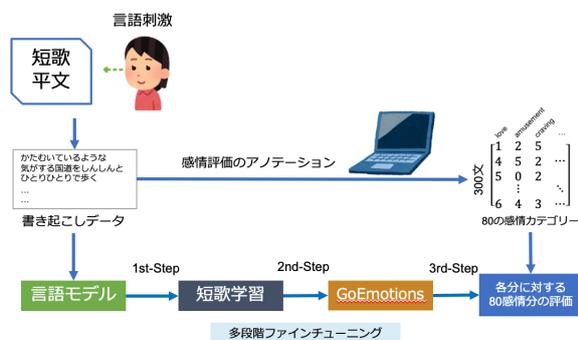


図 1: 研究概要

現代短歌とは、伝統的な和歌の形式を受け継ぎながらも、現代の言葉や表現を用いた 31 音（五・七・五・七・七）の定型詩である。従来の短歌が四季や自然、恋愛などを詠むことが多かったのに対し、現代短歌では、日常生活や個人的な心情を直接的に表現することが特徴の一つとなっている。分析のために、短歌が誘発する感情を識別する感情識別モデルを構築する。感情識別モデルに対してファインチューニングを施し、短歌から抽出される感情の予測精度を向上させることを試みる。また、複数のファインチューニング手法を適用し、それぞれの手法による精度の違いを比較・検証する。学習プロセスでは、モデルの予測ラベルが実測ラベルに近づくように最適化を行い、より精度の高い感情分類を実現することを目指す。

3 感情識別用データ作成

3.1 データ収集

感情収集対象となる短歌

本研究では、短歌が誘発する感情を分析するために、現代短歌と比較用の平文を収集した。短歌データは『現代日本語書き言葉均衡コーパス (BCCWJ)』、『桜前線開架宣言』、および『塔』から抽出した 150 首を用いた。また、比較対象として、BCCWJ に含まれる、短

歌と同じく 31 文字程度の一般的な文 (平文) 150 文を選定した [2, 3].

- 現代短歌の例：
行ってきますと自分に伝えたまいと自分を迎える
単身赴任
- 平文の例：
水がたまった長い坑道を坑内員は休みながら進む

短歌と平文の計 300 文について、言語刺激に曝された際に誘発される 80 の各感情カテゴリーについて、アノテーションを収集した。ここで使用する 80 種類の感情は、Koide-Majima ら [4] による日本語の感情解析に使用された感情カテゴリーに基づいており、付録の表 2 に示す。日本語を母国語とするアノテータを募集し、アノテータは、ある感情カテゴリーが各文から読み取れる感情とどの程度一致しているかを 0 (全く読み取れない) から 6 (非常に強く読み取れる) までの値で評価した。アノテータ自身の感情ではなく、文から読み取れる感情に基づいてアノテーションを行うよう指示した。80 の各感情カテゴリーについて、4 人の異なるアノテータを用意し、独立した 4 つの評価を収集した。80 の感情カテゴリー × 4 つの独立した評価を得るために、合計 384 人 (重複あり) のアノテータをクラウドソーシングサービス・ランサーズ¹ にて募った。分割された 300 文と 5 つの感情の組からなるアノテーション依頼を 1 つずつ出した。既に他の 5 感情でアノテーション作業済みの作業者は同じ文へのアノテーション依頼対象から除く必要があるため、このような依頼の出し方を選択した。80 人のアノテータは各文について、それぞれ割り振られた 5 つの感情カテゴリーに対してアノテーションを行った。このように収集された感情アノテーションは、各感情カテゴリーにて各文ごとに保存された。

3.2 データ前処理

収集したデータには、評価者ごとの感情強度のばらつきが存在するため、データの一貫性を確保するための前処理を行った。まず、各文に対する各感情の感情強度の評価について標準偏差を算出し、ばらつきが閾値を超えるものに関しては、その感情強度をゼロとした。一方、閾値以下の感情強度については、各評価の平均をとることで 24000 件のデータから最終的に 13878 件の統一的なデータを作成した。(実験に使用したデータおよび感情強度のばらつきが観測されたデータの例を付録の図 2, 3 にそれぞれ示す。)

4 実験

4.1 データセット

感情識別モデルの評価のために、BCCWJ に含まれる感情識別用データ作成に未使用の短歌 3571 首を用いた。これらのデータは、マスク言語モデル (MLM)

¹<https://www.lancers.jp/>

における 15% のマスクングを適用したものを使用した。また、比較のために、代表的な感情識別モデルである GoEmotions データセット [5] も活用した。GoEmotions は 27 種類の感情 (+中立) でアノテーションされた英語の感情分類データセットであり、本研究では中立ラベルを除外したデータを用いた。

4.2 実験設定

感情識別には、多言語対応のトランスフォーマーモデルである XLM-RoBERTa の “xlm-roberta-base” の事前訓練済みモデルを使用した。マルチステップファインチューニングを適用し、モデルの性能を比較することを試みた [6]。マルチステップファインチューニングの利点として、まず、異なるデータセットや異なる粒度のラベルを段階的に学習させることで、モデルの汎化性能を向上させられる点が挙げられる。また、初期段階ではより大まかな感情分類を行い、後の段階でより細かい感情分類へと適応させることで、効果的な特徴表現を獲得できる可能性がある。本研究のマルチステップファインチューニングは、以下の 3 段階で構成される。

(1) 短歌の意味理解の学習

BCCWJ に含まれる実験未使用の短歌 3571 首を用いて事前学習を行い、短歌の意味理解を促進する。この段階では、マスク言語モデル (MLM) のタスクを適用し、短歌の特徴を学習させることで、短歌特有の表現を捉えやすくする。

(2) 一般的な感情分類の学習

GoEmotions データセットを用いて、27 種類の感情分類タスクを学習する。このプロセスにより、モデルに一般的な感情識別能力を獲得させる。

(3) 短歌の感情識別の学習

独自に作成した短歌の感情識別用データセットを用いて、短歌に特化した感情強度予測のファインチューニングを行う。この段階では、80 種類の感情を行い、短歌特有の感情表現に適応したモデルを構築する。80 種類という幅広い感情を扱うことで、より細やかな感情表現を反映し、包括的な情動の理解を目指す。

最適化アルゴリズムとして AdamW を使用し、エポック数は訓練サイズに合わせて最適にし、バッチサイズは 64 に設定した。また、単一ステップのファインチューニングや、他の事前訓練済みモデルを用いた場合の精度も評価し、マルチステップファインチューニングとの比較を行った。

4.3 実験結果

実験設定 (1)~(3) のファインチューニング結果を表 1 に示す。精度比較の結果を示す。マイクロ平均の値で比較すると、多段階ファインチューニングをしたモデルの方が F1、精度、再現率ともに高い。

4.4 考察

GoEmotions を基にした 27 種類の感情分類から 80 種類の感情強度の予測へとモデルを拡張した場合は、精度が少し低下した。これは、GoEmotions の学習が英語になっているため、学習が難しくなったと考えら

表 1: 実験設定 (1)~(3) における実験結果

評価指標	(3)	(1)+(3)	(2)+(3)	(1)+(2)+(3)
F1 score	48.26%	48.76%	48.24%	48.78%
Precision	47.75%	46.84%	46.17%	46.52%
recall	48.78%	50.84%	50.51%	51.27%

れる。80 種類の感情強度のみのファインチューニングにおいても、精度の低下はわずかであり、十分に安定した結果を得ることができた。これは、感情強度の予測という新たなタスクに対するモデルの適応力を示唆していると考えられる。

一方で、精度の向上が著しくない点に関しては、感情識別用のデータが十分でなかったことが一因で、さらなるデータがあれば精度向上が期待できると感じた。

5 まとめ

短歌が誘発する情動と特定の感情との関連性を感情という因子を通じて調査した。具体的には、感情識別モデルをファインチューニングし、短歌における感情の予測精度を向上させることを試みた。実験の結果、全ての段階でファインチューニングしたモデルが最も高い精度を示したが、80 の感情カテゴリーのアノテーションのみでの予測においても精度の低下はわずかであり、安定した予測が得られた。この結果から、感情強度の予測タスクにも適応可能であることが示唆された。

一方、感情識別用のデータが十分に収集されていなかったため、今後の研究ではより多くのデータを収集し、モデルの精度向上を目指すべきである。この結果は、短歌における感情表現の特性や、感情を文学的に表現する方法について新たな知見を提供するものであり、今後の研究の発展に寄与する可能性がある。

beginspacing0.95

参考文献

- [1] 三浦衣里, 小林一郎. 言語刺激下のヒト脳内における感情解析への取り組み. In *The 38th annual Conference of the Japanese Society for Artificial Intelligence/ 2024*, 2024.
- [2] 塔短歌会.
- [3] 山田航. 桜前線開架宣言. 左右社, 2015.
- [4] N. Koide-Majima, T. Nakai, and S. Nishimoto. Distinct dimensions of emotion in the human brain and their representation on the cortical surface. *NeuroImage*, Vol. 222, p. 117258, 2020.
- [5] Dorottya Demszky, et al. Goemotions: A dataset of fine-grained emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 4040–4054. Association for Computational Linguistics, 2020.
- [6] Alexis Conneau, et al. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 8440–8451. Association for Computational Linguistics, 2020.

表 2: 80 の感情カテゴリー

love (愛)	joy (喜び)	calmness (穏やかさ)
sadness (悲しみ)	awe (畏怖)	interest (興味)
sexual desire (性的欲求)	tension (緊迫感)	awkwardness (気まずさ)
fear (恐怖)	happiness (幸福)	affection (愛情・好意)
emotional hurt (心の痛み)	empathy (共感)	unrest (動揺)
fever (熱狂)	positive-expectation (前向きな期待)	indecenty (下品)
contempt (侮蔑)	positive-emotion (前向きな感情)	tenderness (優しさ)
relaxedness (リラックス)	negative-emotion (否定的感情)	protectiveness (保護)
cuteness (可愛らしさ)	annoyance (迷惑)	distress (苦痛)
amusement (娯楽)	nostalgia (懐かしさ)	relief (安堵)
admiration (賞賛)	confusion (混乱)	satisfaction (満足)
surprise (驚き)	anger (怒り)	disgust (嫌悪)
horror (bloodcurdling) (恐ろしさ)	friendliness (親しみやすさ)	aggressiveness (攻撃性)
liking (好感)	sympathy (同情)	compassion (思いやり)
exuberance (過剰な喜び)	scare (feel a cill) (ゾットする)	throb (どきどき)
embarrassment (恥ずかしさ)	alertness (警戒心)	vigor (活気)
pensiveness (物思い)	acceptance (容認)	hostility (敵意)
elation (意気揚々)	attachment (愛着)	positive-fear (肯定的恐怖)
stress (ストレス)	craving (渴望)	boredom (退屈)
romance (ロマンス)	aesthetic appreciation (美的満足)	entrancement (魅了、没入)
excitement (興奮)	nervousness (神経質)	anxiety (不安)
empathic pain (共感性苦痛)	laughing (笑い)	ridiculousness (滑稽)
shedding tears (涙を流す)	lethargy (無気力)	curiousness (好奇心)
appreciation of beauty (美の鑑賞)	daze (ぼんやり)	sexiness (セクシーさ)
oddness (奇妙さ)	eeriness (不気味さ)	longing (憧れ)
melancholy (憂鬱)	unease (不安)	levity (平静)
coolness (冷静さ)	encouragement (励まし)	

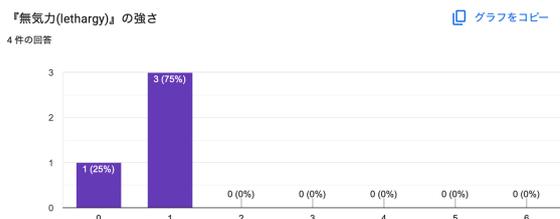


図 2: 閾値以内のデータ (実験使用)

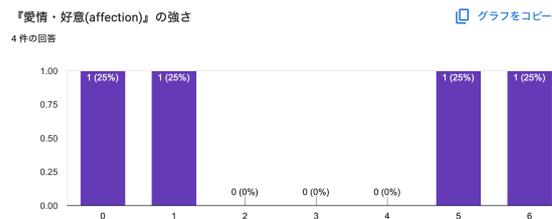


図 3: 閾値超過データ (除外)

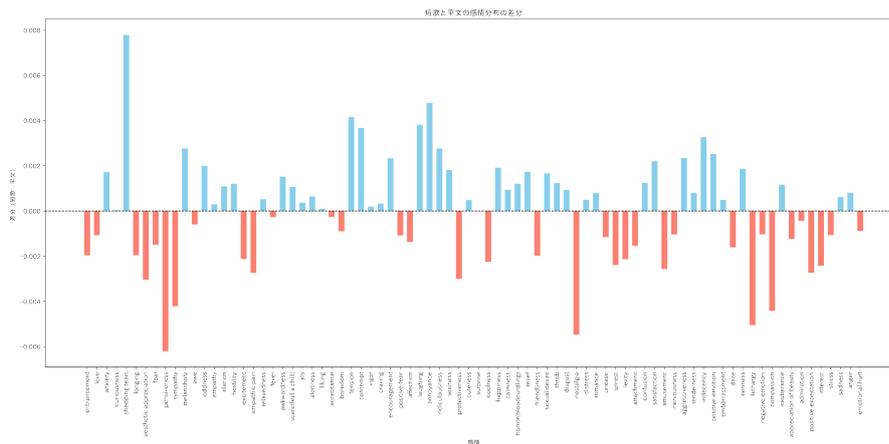


図 4: 短歌と平文の感情分布の差

アンケートの実施

短歌と平文の計 300 文について、80 の感情カテゴリーごとに、各文から読み取れる感情の強度を 0~6 の 7 段階でアノテーションを収集した。評価は、1 つのアンケートにつき短歌・平文各 25 文を提示し、各文について 5 つの感情カテゴリーを対象として行った。1

つの文に対して 4 人の評価データを収集することで、多様な視点を取り入れたデータを構築した。

実験に使用したデータおよび感情強度のばらつきが観測されたデータの例を付録の図 2, 3 にそれぞれ示す。また、短歌と平文の感情分布の差を付録の図 4 に示す。