

SNSにおける混合感情分析と可視化手法

村瀬 礼奈^{1,a)} 馬場 隆寛¹

概要: 本研究では、日本語 SNS 投稿に含まれる混合感情を対象としたマルチラベル分類モデルを構築し、感情の分布構造を分析した。日本語 BERT と双方向 LSTM を組み合わせた分類器を提案し、評価データ 200 件を用いて性能評価を行った。学習率および閾値を変化させた結果、Micro 平均 F1 スコアは最大 0.589 を示し、学習率と学習段階により性能傾向が異なることが確認された。また、BERT 最終層出力を UMAP で可視化したところ、混合感情は極性および覚醒度に沿った分布構造を示し、不安を含む感情は広く分散する傾向が得られた。

今後の課題として、メンタルヘルス・スラングの適切な取り扱い、閾値やラベル付与方針の最適化、個人ごとの投稿履歴や時間的変化を考慮した分析の導入などが挙げられる。

キーワード: 2130102 機械学習, 2130104 ニューラルネットワーク, 2130108 知識発見

Multilabel Classification and Visualization of Mixed Emotions in Japanese SNS Posts Using BERT and BiLSTM

RENA MURASE^{1,a)} TAKAHIRO BABA¹

Abstract: This study investigates multilabel classification and structural visualization of mixed emotions expressed in Japanese SNS posts. We propose a multilabel classifier combining a Japanese BERT encoder with a bidirectional LSTM layer to capture contextual and sequential features of emotional expressions. Performance was evaluated on 200 labeled posts using micro-averaged F1-score, precision, recall, and Hamming loss under different learning rates and threshold settings. The proposed model achieved a maximum micro-F1 score of 0.589, and the results indicate that optimal learning rates vary depending on the training stage. Furthermore, we applied UMAP to the final-layer BERT representations to visualize the distribution of mixed emotions. The embeddings revealed a gradual polarity-arousal structure, where mixed emotions involving anxiety were widely dispersed across the space, suggesting high variability in contextual expression. These findings contribute to understanding and modeling complex emotional states in social media text.

Keywords: Machine Learning, Neural Network, Knowledge Discovery

1. はじめに

近年、若年層の自殺者数の増加が社会問題となっている。厚生労働省の自殺対策白書 [1] によると、自殺者全体としては減少傾向にある一方で、若年層については増加傾向にある。特に、小中高生の自殺者数は、令和 5 年に 513 人となり、

過去最多であった前年と同水準で推移していると報告されている。小中高生の自殺の原因や動機は不詳である割合が高く、家庭問題、健康問題、学校問題などが要因となって命を落とす事例が後を絶たないことが判明している。文部科学省も自殺リスクの把握や適切な支援に努めているが、自殺者数の増加傾向は依然として続いている。

一方で、SNS は若年層にとって主要な感情表現の場として注目されている。劉 [2] らは SNS への帰属意識が高い若者ほど自己開示を行いやすいと示唆したが、鄭 [3] はストレ

¹ 久留米工業大学 工学部 情報ネットワーク工学科
Kurume, Fukuoka 830-0052, Japan

^{a)} e224144mr@kurume-it.ac.jp

ス経験を記録する行為は精神の不健康に関連すると指摘している。

そこで本研究は、SNS 内の投稿に含まれる複数の感情に着目し、それらを混合感情として捉え、機械学習を用いて抽出・分類する手法を提案する。これにより、若年層の感情の複雑さを把握することで、早期の精神的支援により自殺を未然に防ぐための一助となることを期待する。

2. 目的

以上の背景から、SNS に投稿された文面から複数の感情が表出するものに着目し、それらを混合感情として捉え、機械学習を用いて抽出・分類する手法の有効性を検証することを目的とする。そのために、混合感情を捉えることに適した分類モデルについて調査を行う。若年層の SNS 上における感情の複雑さを詳細に把握することで将来的に的確な支援につなげることを目指す。

また、本研究では、作成した分類モデルをもとに、分析結果の可視化を行う。これにより混合感情の分布傾向や感情同士の関係性を明らかにする。

3. 関連研究

SNS 投稿は日常生活における感情表現の場として活用されるとともに、個人の心理状態や社会的動向を反映した、時間的連続性や状況依存性を有するデータとして注目されている。テキストデータを機械学習により解析することで、従来のアンケート調査や面接において生じやすい回答者の主観的バイアスや、調査時点の一時的な心理状態の影響を受けにくく、より日常のかつ継続的な心理状態を捉えられる可能性がある。実際、SNS 投稿を用いた感情推定の研究は多く報告されている。

De Choudhury[4] は、英語圏の SNS 投稿を対象に、投稿内容の言語的特徴と投稿頻度、時間帯、交流のパターンなどの投稿行動を組み合わせることで、利用者の抑うつ傾向を予測する研究を行った。一方、Kitaoka[5] は日本語圏の SNS 投稿を対象に、投稿内容の感情分類を行い、その分布傾向を分析した。その中で、曜日や時間帯、特定の場所における感情の変化について考察し、感情分析の有用性と重要性を示している。

複数の感情が同時に付与される問題に対しては、従来の単一ラベル分類では、投稿に含まれる複数の感情を正確に捉えられない場合がある。こうした背景を踏まえ、複雑な感情を分類するためにはマルチラベル分類の枠組みが有効であるとされている。Zhang[6] は、複数の感情が同時に付与されるタスクにおいて、マルチラベル分類が適切であることを整理しており、各感情ラベルを独立に予測することで、感情の組み合わせや関係性を捉えることが可能であると述べている。

近年の研究では、混合感情を分類する試みが報告されて

いる。味噌岡 [7] らは、単一ラベル分類における抽出精度には限界があると指摘し、マルチラベル分類の有用性を示した。そこで、文中の異なる対象や側面（アスペクト）ごとに感情を推定するアスペクトベース感情分析を提案した。本手法では、深層言語モデルにより文脈情報を学習し、各アスペクトに独立して感情ラベルを割り当てることで、複雑な感情表現や感情間の相互関係を把握可能にしている。評価実験では、従来手法と比較して精度や F1 スコアの向上が確認され、複数感情の同時推定における有効性を示した。さらに、複数投稿をまとめて分析する研究も行われている。Mayor[8] らは、SNS 投稿の時間帯に着目し、個々の投稿を独立に扱うのではなく、一定期間の投稿全体から感情傾向を推定する手法を提示した。これにより、時間帯や曜日による感情の変化、話題転換に伴う感情の動きといった短期的・中期的な変動を滑らかに捉えることが可能となった。利用者の心理状態や感情の傾向をより正確に把握できることを示している。一方、鈴木 [9] らは感情ラベルの付与には個人差が存在することを指摘した。同じ投稿であっても、書き手や読み手の性格、経験、文化的背景、投稿スタイルなどによって感情の受け取り方や表現が異なるため、感情ラベルの一義的な決定が難しいことが示されている。また、丹羽 [10] らも、読み手は書き手の状況を自らの経験や嗜好に基づいて推定し、その心理状態を再現しているため、感情分類における曖昧性や主観性は大きな課題であると指摘した。これらの先行研究は、機械学習モデルでラベルを学習させるだけでは、利用者の感情の多様性や複雑性を十分に捉えられない可能性があることを示唆している。

マルチラベル分類モデルの設計に関して、Devlin[11] らは、Transformer を用いた BERT を提案した。BERT は大規模なデータセットにおいて、文脈情報を深く捉えることが可能であり、長文や複雑な文章構造における意味関係の理解に優れたモデルである。一方、Ezen-Can[12] らは、小規模なデータセットや短文に対しては LSTM が有効であることを示し、データ規模や特性に応じたモデル選択の重要性を指摘した。また、混合感情の分布や感情間の関係性を把握する手法として、Maruyama[13] らは SNS 投稿を対象に、埋め込み表現において感情情報がどの程度保持されるかを分析し、その結果を UMAP による可視化を用いて報告している。UMAP は、高次元の文脈表現を低次元空間に投影することで、感情表現の分布や類似性を直感的に把握できる手法である。また、局所的なデータ構造を保ちながら次元削減を行う特性を有しており、類似した感情を有するデータが近接して配置される。そのため、複雑な混合感情の分布傾向や感情間の関係性を視覚的に分析する手法として有用である。

これらの先行研究の多くは、感情ラベルの推定精度や可視化結果の分析それぞれを主に扱っており、混合感情がどのような分布構造を持つかについて検討している研究は限

られている。この課題を踏まえ、本研究では、日本語 BERT による事前学習済みの文脈表現を入力とし、双方向 LSTM で補完を行うマルチラベル分類モデルを採用することで、大規模データでの文脈理解能力と小規模データにおける学習安定性を両立するモデルを構築する。さらに、構築したモデルを用いて、SNS 投稿に含まれる混合感情を抽出し、抽出結果を UMAP により可視化し、感情の分布や関係性の分析を行う。

4. 提案手法

本章では、SNS 投稿に含まれる混合感情を抽出・分類し、その分布構造を分析するために本研究で用いた手法について述べる。実験では、事前学習済みの日本語 BERT と双方向 LSTM を組み合わせたマルチラベル分類モデルを用いて混合感情の推定を行った。

使用したデータの概要、感情ラベルの設計、モデル構成および学習方法の詳細に加え、推定結果に基づく混合感情の分布構造を可視化する手法について説明する。

5. 実験方法

5.1 データセットの作成

データセットとして、本研究では SNS プラットフォーム X に投稿された文章を対象に、本研究の著者のアカウントを用いて協力者を募り、データ収集を実施した。収集対象は非公開アカウントの投稿とし、研究目的を事前に説明したうえで同意を得た投稿のみを使用している。取得した投稿数は 1000 件である。また、個人が特定される情報については取得せず、研究倫理に配慮した形でデータを取り扱った。

これらの投稿に対して、絵文字や URL など分析に不要な要素の削除、表記の揺れや文体の整形といった前処理を行った。この前処理により、SNS 特有の表現や投稿内容の意味を可能な限り保持しつつ、感情分析に影響を与えるノイズの除去を行った。

感情ラベルは、Joy (喜び)、Anger (怒り)、Sadness (悲しみ)、Anxiety (不安)、Relief (安心)、Surprise (驚き) の 6 種類を設定し、各投稿に対して、文脈全体を考慮したうえで主要な 2 種類の感情を手手で判定し、感情ラベルを付与した。このように、1 つの投稿に対して複数の感情ラベルを付与することで、単一感情では捉えきれない SNS 投稿における混合感情を表現可能とした。

収集したデータの特徴を視覚的に把握するため、いくつかのグラフを作成した。図 1 に各感情ラベルの出現頻度を示す。さらに、図 2 に各投稿文の長さの分布を示す。これらの図から、SNS 投稿データには複数の感情が同時に現れる傾向があり、投稿文の長さにはばらつきが存在することが確認できる。

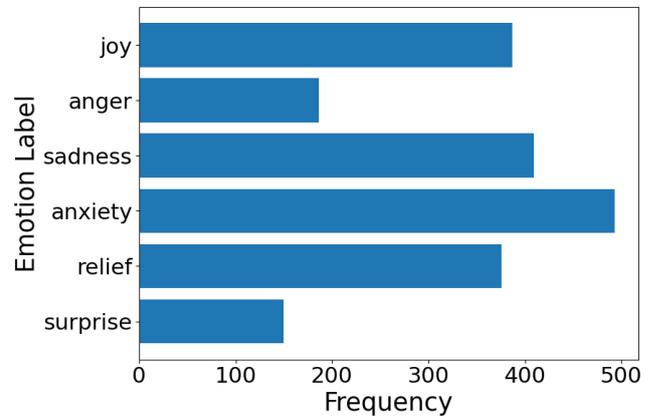


図 1: 各感情ラベルの出現頻度

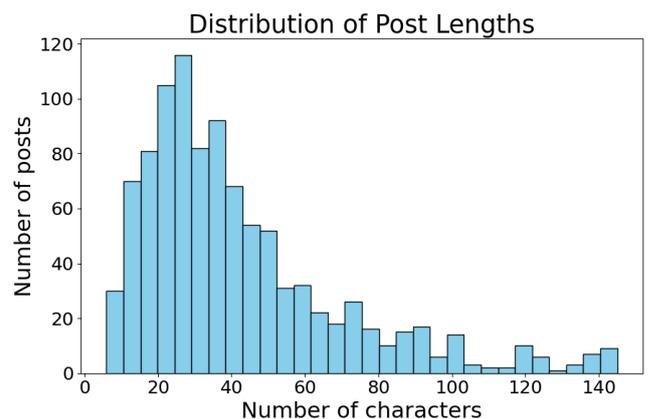


図 2: 投稿文の長さ

6. モデル構築

本研究では、事前学習済みの日本語 BERT と双方向 LSTM を組み合わせたマルチラベル分類モデルを構築した。本モデル構築の狙いは、文脈情報の学習と系列情報の補完を同時に行い、投稿の長さや文章構造に依存せず、複数の感情を正確に抽出することである。

作成したデータセットをもとに、ランダム抽出によりデータの分割を訓練データ 800 件、評価データ 200 件に分割した。また、分割後のデータについては表 1 に示す。

表 1: データ分割

データ	件数	分割割合
訓練データ	800 件	80%
評価データ	200 件	20%

学習設定

モデルの学習にあたり、以下の設定を適用した。

- **トークナイズ:**
事前学習済み日本語 BERT モデル `cl-tohoku/bert-base-japanese` のトークナイザを使用し、入力文をトークン列に変換した。
- **モデル構成:**
BERT の最終隠れ層 (768 次元) を入力として、中間層として双方向 LSTM (BiLSTM) を配置した。
LSTM の隠れ状態次元は 256 とし、双方向化による 512 次元の文表現を得た後、全結合層により各感情ラベルのスコアを計算した。
- **損失関数:**
各ラベルを独立に判定する多ラベル分類タスクであるため、`BCEWithLogitsLoss` を使用した。
- **最適化手法:**
Adam を使用し、学習率を 2×10^{-5} および 3×10^{-5} に設定した。
また、本データセットは規模が小さいため、 1×10^{-5} のような低い学習率ではモデルの更新量が不足し、十分なタスク適応が行われないと判断した。
- **ハイパーパラメータ:**
 - エポック数: 3~10
 - ドロップアウト率: 0.1
 - バッチサイズ: 16
 - LSTM 中間ノード数 (隠れ状態次元): 256
- **出力と二値化:**
モデルの最終出力にシグモイド関数を適用し、各ラベ

表 2: エポック数 7 および 9 における評価結果

Epoch	学習率	閾値	F1	適合率	再現率	HL
7	2×10^{-5}	0.5	0.539	0.577	0.505	0.288
7	2×10^{-5}	0.6	0.522	0.631	0.445	0.272
7	3×10^{-5}	0.5	0.581	0.630	0.540	0.259
7	3×10^{-5}	0.6	0.539	0.636	0.468	0.267
9	2×10^{-5}	0.5	0.589	0.614	0.565	0.263
9	2×10^{-5}	0.6	0.515	0.630	0.435	0.273
9	3×10^{-5}	0.5	0.552	0.598	0.513	0.278
9	3×10^{-5}	0.6	0.548	0.625	0.488	0.268

ルに $[0, 1]$ の確率値を得る。その後、閾値 0.5 および 0.6 に基づいて二値化を行う。

● 評価指標:

マルチラベル分類における性能を比較するため、Micro 平均の F1 スコア・適合率・再現率・ハミング損失 (HL) を算出した。

以上の条件を組み合わせ、学習率・エポック数・閾値の違いに基づく影響を比較した。

7. 実験結果

本研究で構築した日本語 BERT と双方向 LSTM を組み合わせたマルチラベル分類モデルについて、評価データ 200 件を用いて性能評価を行った。学習済みモデルの出力に対してシグモイド関数を適用し、閾値 0.5 および 0.6 で二値化したうえで、Micro 平均の F1 スコア (F1)、適合率、再現率、ハミング損失 (HL) を算出した。

本研究では学習率 2×10^{-5} および 3×10^{-5} 、エポック数 3~10 の条件を組み合わせ検証を行った。その中で性能が比較的高かったエポック数 7 および 9 の結果を代表例として取り上げ、学習率と閾値の違いによる性能変化を比較する。

エポック数 7 および 9 における評価結果を表 2 に示す。

表 2 より、エポック数 7 では学習率 3×10^{-5} かつ閾値 0.5 の条件が最も高い F1 スコア 0.581 を示した。一方、エポック数 9 では学習率 2×10^{-5} かつ閾値 0.5 の条件が最も高い F1 スコア 0.589 となった。

以上より、学習初期段階では高めの学習率 3×10^{-5} が有利である一方、学習が進むにつれて 2×10^{-5} の方が安定した性能を示す傾向が確認された。また、閾値を 0.6 に設定すると適合率は向上するものの再現率が低下し、検出漏れが増加することが分かった。

8. 考察

以上の実験結果より、本研究で提案した日本語 BERT と双方向 LSTM を組み合わせたマルチラベル分類モデルは、Micro 平均 F1 スコアがおおむね 0.55 前後を示し、SNS 投稿に含まれる混合感情を一定程度抽出可能であることが確認された。単一感情分類と比較して難易度の高い混合感情

分類において、BERT による文脈表現と双方向 LSTM を組み合わせた本研究のモデル構成は有効に機能していると考えられる。特に、文全体の意味を踏まえた推定が可能となっている点が本手法の特徴である。

一方で、ハミング損失が 0.26~0.28 程度で推移していることから、感情ラベルの誤検出や検出漏れが依然として存在することが示唆される。本研究では一投稿につき主要な 2 種類の感情ラベルを人手で付与しており、細かな感情の変化や曖昧な感情表現を十分に反映できていない可能性がある。このようなラベル設計は、マルチラベル分類において誤検出が生じやすいという構造的な課題があることが示唆される。また、感情ラベルの付与には人手による主観が含まれるため、丹羽ら [10] が指摘するように、書き手と読み手の違いによる感情理解の揺れも性能評価に影響していると考えられる。

UMAP による可視化結果から、混合感情は明確に分離したクラスターを形成するのではなく、ネガティブ・ポジティブ傾向 (Valence) および覚醒度 (Arousal) の違いに沿って連続的に分布する様子が確認された。喜びと安心を含む混合感情は比較的まとまった分布を示す一方で、悲しみや不安を含む混合感情は UMAP 空間全体に広く分布しており、同一ラベルであっても文脈の違いによる感情表現のばらつきが大きいことが示唆される。この傾向は、感情が連続的かつ多様に表出する Russell の円環モデルの考え方も整合的である。

このような分布傾向は、Maruyama [13] らが指摘するように、文脈埋め込み表現が感情情報を保持する一方で、混合感情を明確に分離することには限界があることを示している。すなわち、本研究のモデルは感情の相対的な関係性を捉えることは可能であるが、感情境界の曖昧さや文脈依存性の高さといった混合感情特有の課題は依然として残されている。

以上より、本研究の手法は、SNS 投稿に含まれる混合感情の抽出に一定の有効性を示すとともに、混合感情が連続的かつ多様に分布する構造を有することを示した。今後は、ラベル設計や閾値設定の改善、時間的・個人差を考慮した分析を行うことで、精度向上と実社会への応用が期待される。

9. まとめ

本研究では、SNS 投稿に含まれる複数の感情に着目し、これらを混合感情として捉え、機械学習を用いて抽出および分類する手法を提案した。近年、若年層における精神的問題や自殺リスクの増加が社会的課題となる中で、日常的な感情表現の場である SNS 投稿を分析対象とすることは、従来のアンケート調査や面接手法を補完する新たなアプローチとして重要性が高まっている。

そこで本研究では、日本語 BERT と双方向 LSTM を組み合わせたマルチラベル分類モデルを構築し、投稿文の長

さに依存しない混合感情の自動抽出を試みた。評価実験の結果、Micro 平均 F1 スコアはおおむね 0.55 前後を示し、単一感情分類と比較して難易度の高い混合感情分類においても、文脈情報を考慮したモデル構成が一定の有効性を有することを確認した。一方で、ハミング損失は 0.26~0.28 程度で推移しており、感情ラベルの誤検出や検出漏れといった課題が依然として残ることも明らかとなった。これらの結果から、混合感情の抽出は一定程度可能である一方、精度向上にはさらなる工夫が必要であることが示唆された。

また、UMAP による可視化を通じて、混合感情は明確に分離したクラスターを形成するのではなく、ネガティブ・ポジティブ傾向 (Valence) および覚醒度 (Arousal) の違いに沿って、連続的かつ多様に分布する構造を有することが確認された。特に、喜びや安心を含む混合感情は比較的まとまった分布を示す一方で、不安や悲しみを含む混合感情は広範に分布しており、同一の感情ラベルであっても文脈や表現の違いによるばらつきが大きいことが示唆された。この結果は、感情が離散的なカテゴリではなく連続的に変化するという心理学的知見とも整合的であり、混合感情分析の困難さと同時にその重要性を示している。

今後の課題として、分類性能向上のためのデータ拡張および大規模化が挙げられる。本研究では 1000 件の投稿を用い、一投稿につき 2 種類の感情ラベルを付与したが、感情表現の多様性や投稿内容の偏りを十分に反映できていない可能性がある。高須ら [14] が示すように、大規模データを用いた分析はモデルの安定性や汎化性能の向上に寄与すると考えられる。また、松崎 [15] が指摘するメンタルヘルス・スラングを含むスラング表現は多義的であり、既存の言語モデルでは感情を適切に捉えきれない場合があるため、追加学習などによるモデル拡張が求められる。さらに、データおよび学習設定の改善により、 1×10^{-5} などの低学習率においても安定した学習を可能とする、実用的な感情分析モデルへの発展が期待される。加えて、閾値設定やラベル付与方針の最適化、投稿履歴や時間的変化を考慮した分析も今後の重要な課題である。

以上より、本研究の手法は混合感情抽出に一定の有効性を示しており、今後はモデルおよび分析手法の改良を通じて、若年層の精神的支援や早期介入に資する実用的応用への展開が期待される。

参考文献

- [1] 厚生労働省：令和 6 年版自殺対策白書，https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hukushi_kaigo/seikatsuhogo/jisatsu/jisatsuhakusyo2024.html. 2025 年 12 月 12 日アクセス。
- [2] 劉星廷，楠見孝：SNS における帰属意識が自己開示に及ぼす効果，日本心理学会大会発表論文集 日本心理学会第 87 回大会，公益社団法人 日本心理学会，pp. 1B-027 (2023)。

- [3] Zheng, S.: SNS にストレス経験を書き込むことと精神的健康は関連するか?, 人間文化創成科学論叢/お茶の水女子大学大学院 『人間文化創成科学論叢』 編集委員会 編, Vol. 24, pp. 99–106 (2021).
- [4] De Choudhury, M. et al.: Predicting depression via social media, *Proceedings of the international AAAI conference on web and social media*, Vol. 7, No. 1, pp. 128–137 (2013).
- [5] Kitaoka, S. et al.: Emotion prediction and cause analysis considering spatio-temporal distribution, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 23, No. 3, pp. 512–518 (2019).
- [6] Zhang, M.-L. et al.: A review on multi-label learning algorithms, *IEEE transactions on knowledge and data engineering*, Vol. 26, No. 8, pp. 1819–1837 (2013).
- [7] 味岡陽紀, 岡田真, 森直樹: 日本語アスペクトベース感情分析における深層言語モデルによるマルチラベル分類手法の拡張と検証, 人工知能学会全国大会論文集 第 38 回 (2024), 一般社団法人 人工知能学会, pp. 1–4 (2024).
- [8] Mayor, E. et al.: Twitter, time and emotions, *Royal Society open science*, Vol. 8, No. 5, pp. 1–26 (2021).
- [9] 鈴木陽也, 山内洋輝, 梶原智之, 二宮崇, 早志英朗, 中島悠太, 長原一: 書き手の複数投稿を用いた感情分析, 人工知能学会全国大会論文集 第 38 回 (2024), 一般社団法人 人工知能学会, pp. 1–4 (2024).
- [10] 丹羽彩奈, 松田寛: 個人間の感情理解の揺れを考慮した感情分析に向けた試み, 人工知能学会全国大会論文集 第 35 回 (2021), 一般社団法人 人工知能学会, pp. 1–4 (2021).
- [11] Devlin, J. et al.: Bert: Pre-training of deep bidirectional transformers for language understanding, *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186 (2019).
- [12] Ezen-Can, A.: A Comparison of LSTM and BERT for Small Corpus, *arXiv preprint arXiv:2009.05451* (2020).
- [13] Maruyama, O. et al.: Preservation of emotional context in tweet embeddings on social networking sites, *Artificial Life and Robotics*, Vol. 29, No. 4, pp. 486–493 (2024).
- [14] 高須遼, 中村啓信, 岸本泰士郎, 狩野芳伸: 大規模ツイートデータを用いたメンタルヘルス不調者の推測, 人工知能学会全国大会論文集 第 36 回 (2022), 一般社団法人 人工知能学会, pp. 1–4 (2022).
- [15] 松崎良美: “メンタルヘルス・スラング” の利用から考える若者の自己認識, 東洋大学社会学部紀要, Vol. 61, No. 2, pp. 45–62 (2024).