

# コロナ禍前後における Twitter ユーザの性格別感情変化の分析

松本 和幸<sup>1</sup> 喜島 涼太<sup>2</sup> 吉田 稔<sup>1</sup> 北 研二<sup>1</sup>

<sup>1</sup>徳島大学大学院社会産業理工学研究部 <sup>2</sup>徳島大学大学院創成科学研究科

{matumoto, mino, kita}@is.tokushima-u.ac.jp, c612135039@tokushima-u.ac.jp

## 概要

2019 年末から始まったコロナ禍も 3 年が経過し、徐々に行動制限が緩和され、日々、状況は変化しつつある。現在も、生活様式の劇的な変化に伴い心身へのストレスなど人々に大きな影響を与え続けている。本研究は、コロナ禍前後での Twitter 上での発言内容から感情分析を行い、特徴的な表現を抽出して比較分析することにより、人々の発言、思考、行動に起きた変化を分析する。また、ユーザの性格タイプ別での感情の変動について考察する。

## 1 はじめに

新型コロナウイルス感染症の拡大に伴い、2020 年からの長期間の行動制限は、精神的ストレスの増加、離職、収入の激減といった悪循環を引き起こしている。現在、行動制限が緩和され、コロナ禍前の日常が戻りつつあるとはいえ、まだ感染が収束したわけではない。現在でも、自粛の継続を選択せざるを得ない人々や、後遺症などから感染拡大前の生活に戻れない人々も少なくない。うつ病などのメンタルヘルスの問題に至らないよう、周囲とのコミュニケーションを密にし、孤立化を防止することが極めて重要である。そのためにも、人々が何に影響を受け、どのような感情を抱いているかを大まかに把握し、要因別のサポート体制を整える必要がある。

本研究では、感情推定技術を活用して、ソーシャルメディア上の発言を感情分析し、人々の発言の内容から感情変化の要因となるものを特定する技術の開発を行っている。人々の感情変化がどのような行動を引き起こしているのか、思考や行動に関連したキーワードを抽出して分析を行い、ユーザを性格別にグループ分けし、どのような性格のユーザにどのような変化が起こっているかを考察する。

## 2 関連研究

コロナ禍における人々の考えや行動をソーシャルネットワークサービス(SNS)の投稿から把握しようとする研究は多い。自粛生活で対面交流が制限され、オンライン交流および Web 上での活動が増加したことで、コロナ禍前よりも多くのユーザが SNS を利用するようになった。Twitter は即時性が高く、短文中で個人の考えをリアルタイムに反映した投稿がされやすい点で、感情分析の対象に適している。また、新しい生活様式においては特に世の中のニュースのチェックに利用する傾向が顕著である[1]。

鳥海ら[2]は、新型コロナウイルスに関する投稿を収集し、情報を発信するユーザの感情に着目した分析を行った。この研究では、新型コロナウイルスに関連した大きなイベントの発生とその際のユーザの投稿に表れている感情との関連を明らかにしている。

福田ら[3]は、新型コロナウイルスワクチンに対する人々の感情に着目し、感情が表れる要因を分析している。日本、米国など 6 カ国（日本語および英語）の Twitter の投稿を対象としており、ワクチンに関する同一の出来事でも、個人の状況に応じて観察される感情が異なることを明らかにした。彼らの研究では、感情分析に BERT を利用し、海外ツイートに対しては SemEval2018 Task 1: Affect in Tweets[4]で配布された Emotion Classification (E-c) task の英語感情分析データセットを用い、日本語ツイートに対しては、Kajiwara ら[5]の日本語感情分析データセット (WRIME corpus)を用いている。

これらの研究では、ユーザ個々の性質には着目していない。鳥海らの研究では、感情の時間変化やユーザのコミュニティの偏りを考慮しているが、個々のユーザの思考・行動の違いは考慮していない。これは、ユーザの思考・行動を識別することが難しいことと、プロフィール情報の信頼性がユーザによって差があること、また、複数アカウントを利用するユーザの影響を排除しきれないことなどが要因であ

ると考える。また、感情分析に辞書ベースの ML-Ask[6]を利用しており精度の面で問題がある。

本研究では、ユーザの感情や行動に影響を与えると考えられる性格（パーソナリティ）に着目し、性格タイプ別に感情の変化を分析することで、思考・行動の変化を検討する点で異なる。

伊藤ら[7]は、IBM のサービス Personality Insights を用いて、Twitter ユーザの性格を時系列で分析している。しかし、個人の性格はコロナ禍前後の短期間で変化が観察できるほど変動の激しいものではない。このため、本研究では、Twitter ユーザ本人が受けた性格診断テスト（MBTI: Myers-Briggs Type Indicator[8]）の結果をもとに性格を固定し分析する。

### 3 提案手法

#### 3.1 感情推定手法

本研究では、感情推定に機械学習アルゴリズムを用いる。具体的には、感情ラベル付きテキストコーパスを、BERT などの事前学習済みの汎用言語モデルを用いて学習させたマルチラベル分類モデルである。学習対象のコーパスは、先行研究と同様に、Kajiwara ら[5]の日本語感情分析データセットと、我々が独自に収集・ラベル付けした感情コーパス[9]、また、絵文字と感情表現辞書をもとにラベルを自動付与したツイートコーパスを用いた。下記の事前学習済みモデルをもとに感情推定モデルのファインチューニングを行う。いずれも Hugging Face[10]で提供されているものである。

1. xlm-roberta-base
2. nlp-waseda/roberta-base-japanese
3. rinna/japanese-roberta-base
4. hajime9652/xlnet-japanese
5. Twitter/twhin-bert-base

ファインチューニングには、Transformer ベースのモデルを簡易な実装で扱えるライブラリ Simple Transformers[11]の MultiLabelClassificationModel を用いた。対象とする感情の種類は、Plutchik の基本 8 感情 [12][13]（喜び:Joy, 悲しみ:Sadness, 期待:Anticipation, 驚き:Surprise, 怒り:Anger, 恐れ:Fear, 嫌悪:Disgust, 信頼:Trust）である。各ツイートに対して推定モデルが出力した確率値が 0.5 以上となった

感情ラベルを付与する。

#### 3.2 感情推定モデルの選定

複数の事前学習済みモデルを用いるにあたり、精度評価を行うことで最も感情分析に適しているモデルを選定する。精度評価には、前節で述べた複数の感情ラベル付きテキストデータから、学習データ 96,696 件、テストデータ 32,233 件を用いた。図 1 に、5 種類の事前学習済みモデルを用いてファインチューニングした感情推定モデルの感情別の F 値（F1-Score）を示す。この図より、XLM RoBERTa が全体的に精度が高いことが分かるため、本研究ではこのモデルを用いて感情分析を行う。

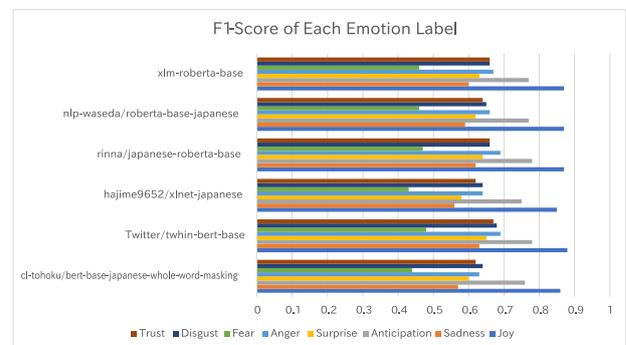


図 1 各モデルの感情ラベルごとの F 値

#### 3.3 性格別感情分析

##### 3.3.1 ユーザアカウントの選定

本研究では、性格診断サイト（16Personalities[14]）を利用して MBTI 診断結果の性格タイプをコロナ禍前後の 2019 年～2022 年の間にツイートしたユーザを選定した。MBTI は、外向型・内向型、感覚型・直感型、思考型・感情型、判断型・認知型の 4 つの二分法を掛け合わせた 16 の性格に分類する自己申告型診断テストの一つであり、大まかな性格タイプを知るために利用される。ユーザが複数回、診断結果をツイートしている場合は、最新の結果をそのユーザの性格とする。性格タイプを、表 1 に示す。

表 1 性格 (MBTI) タイプ

INTJ 建築家	INFJ 提唱者	ISTJ 管理者	ISTP 巨匠
INTP 論理学者	INFP 仲介者	ISFJ 擁護者	ISFP 冒険家
ENTJ 指揮官	ENFJ 主人公	ESTJ 幹部	ESTP 起業家
ENTP 討論者	ENFP 広報運動家	ESFJ 領事官	ESFP エンターティナー

### 3.3.2 ツイートの収集

3.3.1 節で述べた方法で選定した計 2720 名のユーザーごとに、2019 年～2022 年までの期間のツイートを Twitter API を使用して収集した。本研究では、コロナ禍前後を明確に区別することが難しいと考え、2019 年 11 月 30 日以前（新型コロナウイルスが知られる前）をコロナ禍前、約 1 年後の 2020 年 11 月 30 日以降をコロナ禍後と定義する。過去のツイートの収集に、Academic Research 用の API[15]を利用した。性格タイプ別のユーザー数を表 2 に、収集したコロナ禍前後の対象ユーザーのツイートおよびリプライ数を表 3 に示す。リツイートは収集対象から除外した。表 2 の both は、コロナ禍前後両方でツイートを収集できたユーザー、before はコロナ禍前のみ、after はコロナ禍後のみツイートを収集できたユーザー数を示す。

表 2 MBTI タイプごとのユーザー数内訳

MBTI	both	before	after	MBTI	both	before	after
INTJ	160	14	12	INFJ	216	20	23
INTP	123	12	16	INFP	291	34	39
ENTJ	126	13	26	ENFJ	93	6	9
ENTP	92	9	3	ENFP	116	15	14
ISTJ	88	5	3	ISTP	134	9	19
ISFJ	213	20	36	ISFP	183	16	21
ESTJ	110	6	10	ESTP	95	3	12
ESFJ	94	9	12	ESFP	112	11	17

表 3 ツイート数およびリプライ数

	ツイート数	リプライ数
コロナ禍前	149,740	71,051
コロナ禍後	159,715	61,716

### 3.3.3 コロナ禍前後での性格別感情分析

ユーザー集合を性格タイプ別に分割し、ツイート集合を作成する。コロナ禍前後両方でツイートを収集できたユーザーのみを対象とする。性格タイプによるユーザー数の偏りを考慮し、最もユーザー数の少ない性格タイプ ISTJ に合わせて各タイプごとに 88 ユーザーをランダムに選択する。対象ユーザーのツイート集合をコロナ禍前後で分けて、3.2 で学習した感情推定モデルによりマルチラベル感情推定を行う。推定された感情ラベルの頻度を正規化し、コロナ禍前後での性格タイプごとの感情分布の変化を比較分析する。

## 3.4 特徴語分析

ツイートに含まれる単語を、TF-IDF に基づく特徴語抽出を用いて分析する。分析結果から、コロナ禍前後で、どのような変化が起こっているかを考察す

る。特徴語抽出に際して、前処理としてツイートの正規化を neologdn[16]で行い、MeCab で形態素解析し、原形に変換する。ストップワードリストによりノイズとなる語を除去後、名詞・形容詞・動詞・形容動詞語幹などの自立語および固有名詞に限定して TF-IDF 値を計算し、上位 100 語の特徴語を抽出した。

## 3.5 感情変動パターンの分析

ユーザーの感情変動パターンの傾向をコロナ禍前後で分析する。ユーザーごとに、ツイートの感情推定結果をツイートのタイムスタンプの順で並べ（日時の間隔は考慮しない）、感情の種類別に、出現の有無を 0/1 で表し、順番に並べたものを感情変動パターンとする。得られたすべての感情変動パターンを対象に時系列パターンクラスタリング手法 k-Shape[17]を用いて K 個のクラスターに分割する。k-Shape の学習には、Python 用パッケージ tslearn[18]を用いる。各感情ごとの所属クラスターの頻度ベクトルをそれぞれ頻度の総和で正規化して水平連結し、性格タイプごとにまとめる。これを感情変動パターンのクラスター分布ベクトル ( $K \times 8$  次元) とし、コロナ禍前後でユークリッド距離を求め、大きな変化がみられる性格の傾向を考察する。

## 4 分析結果と考察

### 4.1 感情分布の結果

コロナ禍前後での感情推定結果の分布の差分を図 2 に示す。各セルにおける正/負の値は、それぞれコロナ禍後に各感情に推定されるツイートの割合が増加/減少したことを示している。

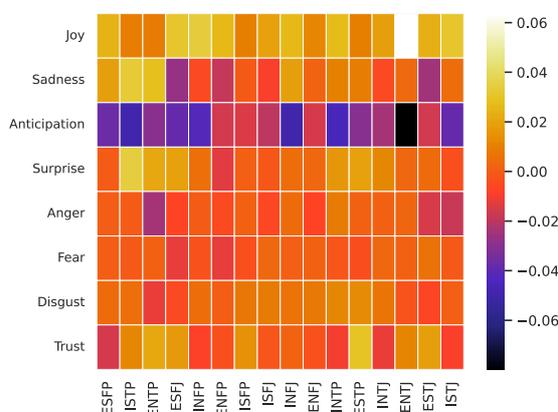


図 2 コロナ禍前後での感情分布の差分

これより、全体的に Anticipation が減少していることが分かる。特に、ENTJ（指揮官）については、Anticipation の減少、Joy の増加が目立つ。また、感情分布ベクトルの Kullback-Leibler divergence が最も大きくなったのは、ENTJ（指揮官）、次いで ENTP（討論者）、最も小さくなったのが ENFJ（主人公）であった。

## 4.2 特徴語分析の結果

感情分布の変化が大きい ENTJ と ENTP に関して、コロナ禍前後で共通している語、コロナ禍前のみに出現した特徴語、コロナ禍後のみに出現した特徴語の例を表 4 に示す。ENTJ では、コロナ禍後に「おめでとう」「楽しみだ」などのポジティブ表現、また、「測定」「体重」など健康を気遣うような表現がある。一方、ENTP では、コロナ禍後に「最高だ」「面白い」などポジティブ表現が増加している。

表 4 コロナ禍前後での ENTJ と ENTP の特徴語分布

	ENTJ	ENTP
共通	仕事, 笑う, ありがとう, 買う, くれる, 好きだ, 食べる, 日本, 募集, 本当だ, etc.	笑う, ありがとう, 好きだ, 仕事, 欲しい, 楽しい, 食べる, 本当だ, etc.
前	考古, ドッカンバトル, ドラゴンボール, ヤング, ガルフレ, 世界, 願う, イベント, 公式, 大阪, etc.	願う, 一緒だ, 問い合わせ, 東京, TEL, Minecraft, ポケモン, 予約, 結果, 当たる, 飲む, etc.
後	結果, 配信, 観る, 簿記, おめでとう, 変わる, 筋肉, もらう, 必要だ, 楽しみだ, 測定, 体重, etc.	読む, 記念, 最高だ, かわいい, 面白い, 寝る, すごい, 世界, 頑張る, 遊ぶ, マジ, 嬉しい, etc.

## 4.3 感情変動パターン分析の結果

クラスタ数  $K$  を 4 に設定し、時系列クラスタリングを行う。各クラスタへの所属回数をカウントして正規化したベクトル間のユークリッド距離を、コロナ禍前とコロナ禍後で比較したものが、図 3 である。“euclid\_before/after”がコロナ禍前後での同一性格タイプ間の距離、“euclid\_same\_avg”がコロナ禍前後それぞれの同一期間における他の性格タイプとの平均距離、“euclid\_before\_avg”がコロナ禍前の他の性格タイプとの平均距離、“euclid\_after\_avg”がコロナ禍後の他の性格タイプとの平均距離を示している。この図より、ISFP（冒険家）、ESTJ（幹部）、ESFP（広報運動家）の 3 つのタイプは、平均距離よりも距離が大きく、感情変動パターンの変化が比較的大きい。一方、コロナ禍前後の距離が最も小さかったのは INFJ（提唱者）であった。

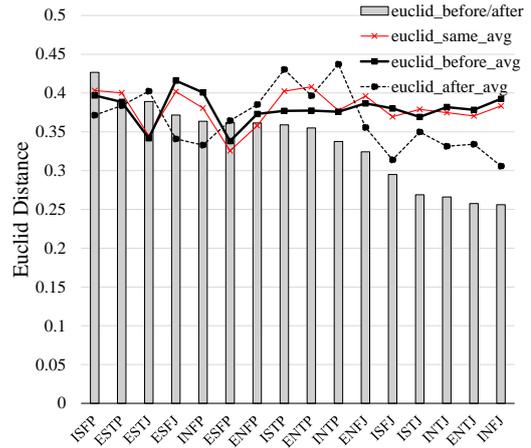


図 3 コロナ禍前後の感情変動パターンの距離比較

## 4.4 考察

コロナ禍前後での特徴語の変化を分析した結果から、特定の性格タイプにおいては、発言に特徴的な変化がみられた。その結果、発言から推定できる感情にも変化が起きたと考えられる。また、コロナ禍前後での感情変動に変化が見られたのは、型破りなタイプである ISFP（冒険家）や、リーダーシップをとるタイプである ESTJ（幹部）などであった。感情変動に変化があるということは、思考や行動にも変化が起きている可能性があるため、特徴語やトピックの分析をすることで確認する必要がある。

## 5 おわりに

本研究では、コロナ禍前後のツイートの感情推定を行い、性格タイプ別の分析を行った。感情分布や感情変動には、ある一定の変化がみられたが、それらの要因まではとらえきれていない。感情分布の変化が大きい性格タイプのユーザの特徴語を分析したところ、コロナ禍後に変化が起こっていることは確認ができたが、現時点では性格との関連性については明らかにできていない。また、感情によって変動パターンのバリエーションが豊富な場合も考えられるため、クラスタ数を感情ごとに最適化する必要がある。今後は、コロナ禍後のデータをより長期的に収集し、感情分析とトピック分析を併用し、特徴的なイベント（緊急事態宣言など）の発生と感情および性格タイプとの関係を明らかにしたい。

謝辞

本研究は JSPS 科研費 JP20K12027, JP 21K12141, SCAT 研究助成の助成を受けたものです。

## 参考文献

- [1] 新型コロナがもたらした【新しい生活様式】における消費者の SNS 利用実態調査. 日本語組版の要件 (日本語版). echoes : <https://service.aainc.co.jp/product/echoes/voices/0033>, 2023-1 閲覧.
- [2] 鳥海 不二夫, 榊剛史, 吉田光男. ソーシャルメディアを用いた新型コロナ禍における感情変化の分析, 人工知能学会論文誌, Vol.35, No. 4, pp. F-K45\_1-7, 2020.
- [3] 福田 悟志, 難波英嗣, 庄司裕子. コロナ禍におけるワクチンに対する人々の感情変化とその要因の分析, 知能と情報, Vol.34, No.3, pp. 592-600, 2022.
- [4] S. Mohammad, et al. SemEval-2018 Task 1: Affect in Tweets. Proceedings of the 12th International Workshop on Semantic Evaluation, pp.1-17, 2018.
- [5] T. Kajiwara et al. WRIME: A New Dataset for Emotional Intensity Estimation with Subjective and Objective Annotations. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp.2095-2104, 2021.
- [6] M. Ptaszynski, et al. ML-Ask: Open Source Affect Analysis Software for Textual Input in Japanese, Journal of Open Research Software, Vol.5, No.1, 16, 2017.
- [7] 伊藤 桃, 榎 美紀, 小口正人. ソーシャルメディアを利用した新型コロナ禍におけるスポット別性格値推移に関する調査. DEIM Forum 2021 I25-2.
- [8] I. B. Myers and P. B. Myers. Gifts Differing: Understanding Personality Types, Davies-Black Publishing, Mountain View, CA, 1995.
- [9] 日本語感情コーパス Tokushima Univ. A-2 Lab.: <https://github.com/Kmatsu-tokudai/emotionCorpusJapaneseTokushimaA2Lab> 2023-1 閲覧.
- [10] Hugging Face: <https://huggingface.co/> 2023-1 閲覧.
- [11] Simple Transformers: <https://simpletransformers.ai/> 2023-1 閲覧.
- [12] R. Plutchik. The Multifactor-Analytic Theory of Emotion. Psychology Vol. 50, pp. 153-171, 1960.
- [13] R. Plutchik. The nature of emotions. American Scientist, Vol. 89, Iss. 4; pp. 344-350, 2001.
- [14] 16Personalities: <https://www.16personalities.com/> 2023-1 閲覧.
- [15] Twitter API Academic Research アクセス: <https://developer.twitter.com/ja/products/twitter-api/academic-research> 2023-1 閲覧.
- [16] neologdn: <https://github.com/ikegami-yukino/neologdn> 2023-1 閲覧.
- [17] Zachary G. Ives. Technical Perspective: k-Shape: Efficient and Accurate Clustering of Time Series. SIGMOD Record, March 2016, Vol.45, No.1.
- [18] tslearn: <https://tslearn.readthedocs.io/> 2023-1 閲覧.