

# Twitter 上に出現する形容詞の分析による 日本のインフルエンサーの人気要因の推定

増田 俊太郎 山崎 俊彦

東京大学大学院情報理工学系研究科 〒 113-8656 東京都文京区本郷 7-3-1

E-mail: {masuda,yamasaki}@cvm.t.u-tokyo.ac.jp

**あらまし** ソーシャルメディアの発展により、企業や個人が社会から人気を獲得することがますます重要になってきた。人気獲得が激化する社会において、大衆から長期間支持を得るためには、人気獲得のメカニズムをより深く理解することが求められる。本研究では、日本のインフルエンサーを対象とし、彼らに対する定性評価と人気度の関係性を調べることで人気の要因推定を試みた。定性評価を示す情報として、Twitter 上の投稿から対象と共起する形容詞を抽出し、その出現頻度や、感情分析による紐づく感情の分類を行い、人気を表す定量指標との関係性を調べた。結果、形容詞と感情が対象の人気と一定の関係性を持つことが示唆された。

**キーワード** SNS, Twitter, 人気, 感情分析, 自然言語処理, インフルエンサー, 魅力工学

## 1. はじめに

近年の情報メディアの変容・発達により、大衆の関心や注目の度合いが重要な資源として扱われるアテンションエコノミー[1]が社会に深く浸透し、あらゆる組織や個人にとって人気を獲得することの重要性が増してきている。しかし同時に、あらゆるコンテンツ（ヒト・モノ・コト）が人気獲得の競争に参入しており、大衆から長期間にわたって多くの人気を得続けることは容易ではなくなっている。このような社会背景において、各々のコンテンツの人気を向上させるために人気の発生や維持に関するメカニズムを深く理解することが望まれる。

人気に関する分析手法として、オピニオンマイニングや感情分析などにより大衆側の評価・感情・態度を解析するアプローチ[2, 3, 4]や、画像や音声について、コンテンツの特徴量や **Social dynamics** といった社会的要素を元に人気形成の予測や部分的な要因分析を行うアプローチ[5, 6]がある。しかし、これらのアプローチは単なる意見の分類タスクに特化していたり、適用範囲が特定の対象・シチュエーションに限られていたり、コンテンツリリース直後の比較的短期間の人気推移の予測に焦点が当てられていたりして、人気の"要因"にフォーカスした先行研究は限られている。「なぜこのコンテンツが人気なのか」「対象コンテンツの人気を長期的に向上・維持させるには何を伸ばせばよいのか」といった人気獲得のためのクリティカルな問いに答えるためには、人気の要因部分にフォーカスした分析手法が求められる。さらに、先行研究と比べて、特定のモノだけではなく多様なモノ・ヒト・コトに適用できるような高い汎用性を持ち、かつ人気獲得のための具体的なアクションに繋げられるような解釈性の高い手法であることが望ましい。本研究ではこのような要件を

満たした人気要因分析手法を提案する。

上記を踏まえて、本研究では人気の"要因"を推定することにフォーカスするため、人気の決定者である"大衆"の意見を分析インプットとする新たな分析手法を提案する。人気分析に関する先行研究[5, 6]では、前述したように研究対象のセグメントに応じて設定した固定的な要素（対象コンテンツの特徴量や、社会的要素など）を分析インプットとして比較的短期の人気推移予測や部分的な要因分析を行っている。しかし、このアプローチだと設定した要素以外に重要な要因がある場合に特定できない。また、短期ではなく長期間で時々刻々と変化していく人気に対して適用シーンが限られてしまう。それに対して提案手法では、対象に関するあらゆる観点での大衆意見（評価・感情・態度）を任意の時点で収集し、分析インプットとする。これにより、対象セグメントおよび分析の時間軸を限定せず、かつあらゆる観点を分析範囲に含められるようになり、人気の根本的な要因を抽出できるようにすることが狙いである。

本研究では、YouTube 上で活動しているインフルエンサー（YouTuber）51名を分析対象とし、彼らに対して大衆から向けられている意見について分析することで、YouTuber という特定のセグメントにおける人気の要因を探り、提案手法の有用性について検証を行う。大衆からの意見は、日本でメジャーな SNS の 1 つである Twitter のデータから収集する。Twitter 上の対象インフルエンサーの名称を含む Tweet（テキスト文章）を抽出し、それと共起する単語を抽出する。本研究では、単語の中でもオピニオンマイニングにおいて人の意見の中で重要な情報を含む形容詞と感情[2]を対象を絞り、Twitter 上での大衆からの支持を表す指標であるフォロワー数との関係性を調べる。具体的には、各形容詞の出現割合、および形容詞に紐づく感情を分類

器によって判定して各感情の出現割合を求め、各形容詞・感情の出現割合の順位と、フォロワー数との関係を Spearman の順位相関係数[7]で算出する。分析結果として、特定の形容詞・感情がフォロワー数と有意に相関を持っていることが明らかになったため、本稿の後半で詳細を述べる。

本研究における成果をまとめると以下の2点である。

- ・ 大衆の意見（形容詞・感情）を分析インプットとすることで、人気要因推定にフォーカスした汎用性・解釈性の高い分析手法を新たに提案
- ・ YouTuber を対象とした分析の結果、特定の形容詞・感情が、人気を示す指標（Twitter上のフォロワー数）に対して有意に相関を持つことを示唆

本稿では2章で意見や人気について扱った関連研究について紹介し、それを踏まえて3章で提案手法の詳細について説明する。4章では実験に用いるデータセットの内容を、5章で形容詞・感情とフォロワー数の関係性分析に関する実験結果について説明し、6章では実験結果の解釈など本研究について考察を行う。最後に7章で本研究のまとめと今後の展望について述べる。

## 2. 関連研究

本章では本研究の趣旨に照らし、オピニオンマイニング、感情分析、人気分析に関する研究について紹介する。

大衆の意見を分析する研究分野として、オピニオンマイニングと感情分析がある。これら2つは多くのケースで同様の文脈・アプローチで扱われることが多く、人の消費行動やマーケティング、選挙などの政治的活動を分析する文脈で適用されることが多い[2]。感情分析は主に感情をポジティブ・ネガティブの極性に分類するタスクにフォーカスされることが多いが[2]、近年では感情のアスペクト（対象観点）や意見語の抽出、それらのペアの特定や[8]、テキスト以外の情報も考慮するマルチモーダル感情分析[9]等のより高度なタスクへの試みおよび精度向上に関する研究が盛んである。感情分析は個別意見および大衆の意見から感情極性等の意味のある情報を抽出し、詳細な分析をするための手法として優れているが、感情分析のタスク単体で人気要因を特定するような研究はない。しかし、意見や感情の分析上重要な指標となりうる形容詞[2]や、抽出した感情といった情報は、人からの評価によって形成される"人気"の要因を解像度高く分析するための材料となりうる。

続いて、人気の分析に関する関連研究について紹介する。人気の"予測"に関する研究は一定数あるものの[10, 11, 12]、人気の"要因"の分析にフォーカスした先

Table 1 Tweet 収集方法および内容

Tweet収集方法	Twitter API
対象期間	2022年9月1日～2022年11月30日
総件数	2,276,679
Tweet例	「久しぶりにはじめしゃちょーの動画見てるけど面白い一瞬でハマる」 「はじめんかわいすぎる」 ※はじめしゃちょーに関するTweet例

行研究は限られている。参考になる事例として Ren による音楽の人気に関する研究がある[5]。Ren らはオンライン音楽サービスである Last.fm 上の楽曲を対象に、楽曲の特徴（テンポ、メロディなど）、アーティストの評判（受賞歴）、社会的コンテキスト（リリース後5週間のコメント数など）の3つの要素から Random Forest 等の機械学習手法を用いて人気の予測を行った。Ren らの研究は人気予測に主眼を置いているものの、予測に用いる要素を変えた時の予測精度の違いを調べ、3要素それぞれが人気に貢献していることを示した。しかし、人気の"要因"を追究する場合、各要素の詳細な寄与度の測定や、対象分野や時間軸に応じて柔軟に要素設定できるような手法設計が求められる。また、別の事例として Khosla らによる画像の人気に関する研究がある[6]。Khosla らは写真の共有を目的とした SNS である Flickr 上の画像を対象に、どのような要素が画像の人気に寄与するのかを調べた。分析手法として、画像自体の特徴量や、画像の投稿者の Flickr 上での特徴などのソーシャルキュー（社会的合図）を示す特徴量を説明変数、画像の閲覧数を人気を表す目的変数として、Spearman の順位相関係数や、Support Vector Regression (SVR) [13]等の手法を用いて変数間の関係性を調べた。Khosla らの研究は分析対象も分析インプットとして用いている要素も本研究と異なるが、人気を示す指標とそれを説明する他の要素との関係性を調べるためのアプローチとして応用範囲が広く、各要素の人気に対する影響度合いも解釈しやすい。

## 3. 提案手法

本研究では前述の関連研究の考え方、手法を取り入れ、大衆の意見から分析対象である日本で知名度の高いインフルエンサー（YouTuber）の人気のキーファクターを推定する分析手法を提案する。具体的には、Twitter 上の Tweet から対象インフルエンサーに関連するコメントを含むテキストデータを収集し、テキストデータから形容詞を抽出し、それらの形容詞の出現割合と、感情分類器によって判定した各形容詞に紐づく感情の出現割合を説明変数とし、Twitter 上での対象ア

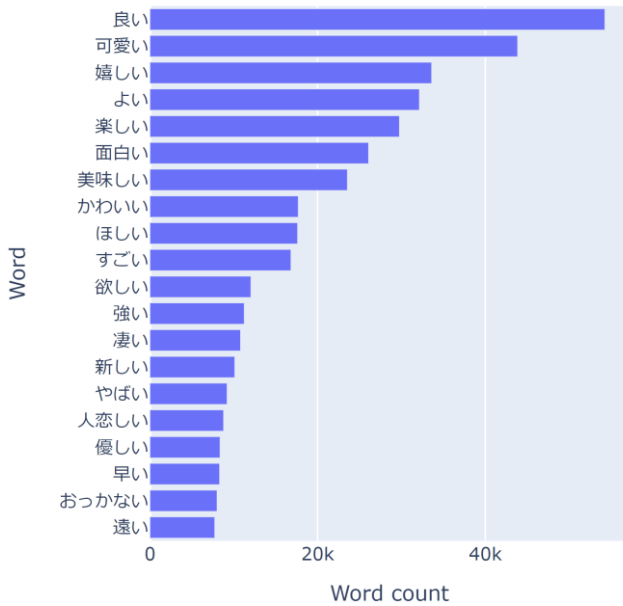


Figure 1 各形容詞の出現頻度

カウントの人気指標の1つであるフォロワー数を目的変数として、これらの変数の関連性を Spearman の順位相関係数によって算出する。

### 3.1. 形容詞の抽出

本研究では Twitter 上の対象インフルエンサーに関するコメントから形容詞を抽出する。対象インフルエンサーとして、日本で Twitter フォロワー数が上位に位置する YouTuber を設定し、彼らを特定する名称（例：ヒカキン）を含む Tweet を収集する。それらの Tweet を Python のライブラリである McCab[14]を用いて形態素解析し、対象名称と共起する形容詞のみを抽出する。なお、McCab による解析の際には、Web 上の言語資源から新語を追加して拡張している McCab 用のシステム辞書 mecab-ipadic-NEologd[15]を利用する。

前述した「共起」とは、ある対象語とセットで特定の語が出現することである。本研究では、同じ Tweet の中に複数文章が含まれた場合、対象インフルエンサーの名称が含まれる文章中に存在する形容詞のみを共起する語として抽出する。この処理により、対象とは関係のない形容詞がノイズとして抽出されてしまうことを抑える。このノイズ低減のための措置としては、岩成ら[16]ら等の研究で用いられている係り受け解析を用いる方が、一般的により精度高く対象に対する形容詞を抽出できる。しかし、係り受け解析を用いる場合、代名詞（例：彼）が含まれる場合等に扱いが難しいことと、最終的に抽出される形容詞の数が大きく減少して分析に支障が出るような懸念から、本研究では前述の共起を基準として形容詞を抽出する。

単純に共起語のみ抽出するとなると、実際には分析対象以外の語に対する形容詞が抽出されるケースが発

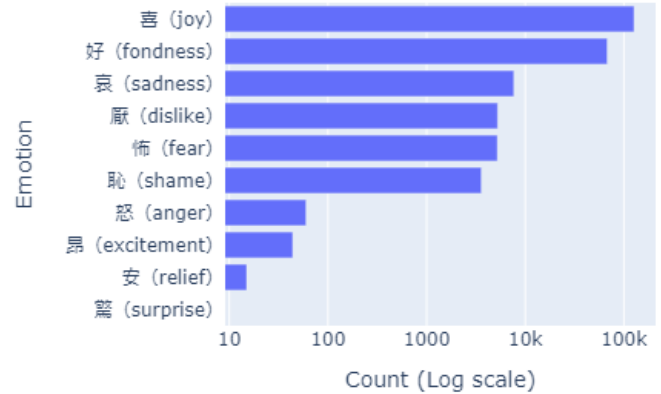


Figure 2 各感情の出現頻度

生ずる。これがデータのノイズとなりうることは免れないが、本研究ではあくまで分析対象と共起する形容詞の全体傾向を捉えることを優先している。

なお、形容詞に否定語が付属している場合、意見としては反対の意味を持つことになるため、この否定の情報を持保持するため、係り受け解析によって形容詞の直後に否定の助動詞が付いている場合は、抽出した形容詞にラベリングして区別している（本稿では否定語が付帯する形容詞には末尾に「(否定)」を付け、「嬉しい(否定)」のように表記する）。

また、「良い」と「よい」、「凄い」と「すごい」など同じ形容詞でも漢字を使われているケースと使われていないケースがあるが、テキストで扱う場合は Tweet の投稿者が表現上意図的に使い分けている可能性もあるため、別々の語として分析を行っている。

### 3.2. 感情分類

抽出した形容詞の感情を判定するため、本研究では日本語の感情多値分類のために用いられる Python ライブラリの ML-Ask[17]を利用する。ML-Ask はインプットテキストを「喜 (joy)」「怒 (anger)」「昂 (excitement)」「哀 (sadness)」「好 (fondness)」「怖 (fear)」「安 (relief)」「厭 (dislike)」「驚 (surprise)」「恥 (shame)」の 10 種の感情に分類するモデルである。このモデルは中村による感情表現辞典[18]をベースに構築されており、キーワードベースでテキストの感情を判定する。このデータベースでは 2100 語の日本語の単語に対して前述の 10 種類の感情がラベリングされている。

ただし、ML-Ask はキーワードベースの分類器であり、対象語がデータベースとマッチしない場合は感情分類がなされない。後述の感情に関わる分析においては、ML-Ask により感情分類がなされた語のみを分析対象としていることに注意されたい。

### 3.3. 順位相関係数

抽出した形容詞、および各形容詞に紐づく感情を判定し、各対象インフルエンサーに関する意見全体の中

**Table 2 形容詞とフォロワー数の順位相関係数  
(上位 10 件)**

形容詞	順位相関係数	利用文脈の事例
美味しい	0.378**	食に関する動画等への感想
古い	0.374**	動画中に紹介された古いコンテンツ等に対する感想
凄い(否定)	0.372**	「凄くない?」と問いかけの形で対象への称賛の気持ちを表現
不味い	0.371**	食に関する動画等への感想
高い	0.366**	(利用される文脈が多岐にわたる形容詞のため割愛)
熱い	0.362**	特定の動画・企画等について関心が高まっている気持ちを表現
旨い	0.356*	食に関する動画等への感想
奥深い	0.339*	対象の動画コンテンツ・スキル・人柄等に関する評価
やばい(否定)	0.311*	「やばくない?」と問いかけの形で対象への驚嘆の気持ちを表現
上手い	0.311*	対象のスキル・パフォーマンス等に関する評価

(否定): 否定語, \*:  $p < 0.05$ , \*\*:  $p < 0.01$ ,

**Table 3 形容詞とフォロワー数の順位相関係数  
(下位 10 件)**

形容詞	順位相関係数	利用文脈の事例
久しい	-0.404**	対象の動画やイベントを久しぶりに見た際の感想
めでたい	-0.354*	対象の誕生日等を祝う気持ちを表現
眠たい	-0.353*	Tweet投稿者自身、もしくは対象の眠気を表現
寂しい(否定)	-0.345*	対象の休暇・活動休止等に対して「寂しくない」と気持ちを表現
暖かい	-0.344*	対象の人柄・発言に関する優しさを評価(「温かい」の誤用)
早い(否定)	-0.332*	(利用される文脈が多岐にわたる形容詞のため割愛)
ねむい	-0.331*	Tweet投稿者自身、もしくは対象の眠気を表現
狡い	-0.328*	対象の発言・行動に対する評価
偉い	-0.327*	対象およびTweet投稿者自身に対する評価
大人っぽい	-0.315*	対象の見た目や動画の内容に対する評価

(否定): 否定語, \*:  $p < 0.05$ , \*\*:  $p < 0.01$ ,

で、各形容詞と感情がどのくらいの割合で出現しているかを算出する。ここでいう出現割合とは、対象インフルエンサーごとに抽出した全形容詞のカウント総数を母数とし、そのうち各形容詞が母数に対してどの程度の比率含まれるかを示す。感情に関しても同様の考え方で出現割合を算出している。

前述の形容詞・感情の出現割合について、形容詞・感情ごとに各インフルエンサー(全 51 名)間で順位付けをし、各インフルエンサーのフォロワー数の順位との相関を算出する。本研究ではこの相関を Spearman の順位相関係数[7]を用いて算出している。2変数  $X$  と  $Y$  の間に  $n$  対のデータがあり、かつ同順位のデータを含まない場合、Spearman の順位相関係数  $r_s$  は以下の式で算出される。

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

$d_i$  は  $n$  対の  $i$  番目のデータ  $X_i$  と  $Y_i$  の順位の差を示しており、全ての順位が一致する場合は  $r_s = 1$ 、全て逆順になる場合は  $r_s = -1$  となる。この相関係数を求めることで、どの形容詞・感情がどの程度人気に関わるかを調べることができる。また、相関係数と合わせて  $P$  値を算出し、その有意性を検証する。

#### 4. データセット

本研究では 2022 年 5 月 30 日～2022 年 11 月 30 日の期間の Twitter データを収集した。データは各対象インフルエンサーに関する Tweet とフォロワー数の 2 種類である。対象インフルエンサーの選定方法も含め詳細を述べる。

#### 4.1. 対象インフルエンサーの選定

本研究では Twitter で上位のフォロワー数を持つ日本のインフルエンサーを対象とした。また、人に関する人気は、属するセグメントによって人気の基準が異なる可能性がある(例: お笑い芸人と俳優など)ので、対象セグメントを YouTuber に絞ることとした。Twitter 上で上位のフォロワー数を持つ日本国内の YouTuber を選定し、上位層 51 アカウント(個人/グループ両方含む)を対象とした。対象の情報は付録 A を参照されたい。

#### 4.2. 収集対象 Tweet

Tweet の収集には Twitter API を利用した。Tweet の抽出条件としては、Tweet の本文内に対象インフルエンサーの名称を含むすべてのものを対象とした。インフルエンサーによっては呼称が複数ある(例: ヒカキンと HIKAKIN など)ため、社会一般的に用いられている呼称を調べて検出キーワードとした。また、一部インフルエンサーは呼称が別人物や別キーワードと被る可能性が高かった(例: 2 人組 YouTuber 「水溜りボンド」のトミーなど)ため、その場合は対象の所属グループ等のキーワードをクエリに AND 条件で加えることで、無関係な Tweet の抽出を抑えるようにした。

Tweet は投稿日時が 2022 年 9 月 1 日～2022 年 11 月 30 日の期間中のものを対象とし、合計約 228 万件の Tweet を取得した(Table 1)。Twitter API の利用上の制限により、対象によっては Tweet の収集期間や件数が異なるが、おおむね数千～数万件程度の Tweet 数になるように収集した。

Table 4 感情とフォロワー数の順位相関係数

感情	順位相関係数
恥 (shame)	0.310*
怒 (anger)	0.254
安 (relief)	0.181
怖 (fear)	0.172
喜 (joy)	0.120
厭 (dislike)	0.054
好 (fondness)	0.040
昂 (excitement)	-0.180
哀 (sadness)	-0.200
驚 (surprise)	-

\* p < 0.05

#### 4.3. フォロワー数

Twitter のフォロワー数は過去断面の数値を取得する必要があったため、Paul らの研究[19]にならい、世界中のウェブページ情報をアーカイブしている「Internet Archive」の「Wayback Machine」(<https://archive.org/web>)という Web サービスを利用した。これにより、各対象の 2022 年 5 月 30 日～2022 年 6 月 8 日の期間中時点のフォロワー数を取得した（詳細なフォロワー数と順位は付録 A を参照）。なお、上述サービスで取得可能なデータに制約があり、4.2 節で述べた Tweet の取得期間とズレがあるが、フォロワー数順位の変動は限定的であり、分析に大きな影響は発生しないと判断した。

### 5. 実験

まず全体の傾向を見るため、形容詞の出現件数について説明する。全対象インフルエンサーの形容詞の出現数（全 673,540 件）の合計値の上位 20 件を Figure 1 に示す。これを見ると、「良い」、「可愛い」、「面白い」などが上位に来ており、日本の人気 YouTuber に関する Tweet の中にはポジティブな形容詞が多く出現していることが分かる。

続いて 10 種の感情の出現件数（全 216,853 件）を Figure 2 に示す。感情によって出現頻度に偏りがあったため、横軸のカウント数を対数表示している。これを見ると、上位から順に「喜 (joy)」、「好 (fondness)」といったポジティブな感情が多く出現しており、続いて「哀 (sadness)」、「厭 (dislike)」などのネガティブな感情が出現していることが分かる。続く節ではこれらの形容詞および形容詞に紐づく感情の出現割合と、フォロワー数の順位との関係性について分析結果を述べていく。

#### 5.1. 形容詞とフォロワー数の順位相関係数

形容詞の出現割合の順位と、インフルエンサーのフォロワー数の順位相関係数の算出結果について述べる。Table 2 と Table 3 に、Spearman の順位相関係数の上位、下位 10 件の形容詞を示している。

まず順位相関係数上位 10 件の Table 2 から見ていくと、「美味しい」「旨い」「不味い」などの食に関する形容詞が多く挙がっていることが分かる。分析対象の中には料理系動画等の食に関するコンテンツをメインで配信しているインフルエンサーは含まれていないが、食に関する世間からの注目が高く、このジャンルの感想の多いほどフォロワー数が多いという傾向が現れている可能性がある。それ以外の目立つ傾向としては、「凄い」や「やばい」の否定語（実際の Twitter 上での主な利用文脈としては、インフルエンサーのパフォーマンス・発言・見た目等の面白さ・奇抜さ・かっこよさに関して「凄くない?」「やばくない?」などと疑問形で賛辞・尊敬・驚嘆の気持ちを表現していることが多い）、「奥深い」「上手い」など、対象インフルエンサーに対する称賛、驚嘆の気持ちを表す形容詞がフォロワー数順位に対して正の相関を持っていることが分かる。上位 10 位でも相関係数が 0.4 未満の値ではあるものの、いずれも P 値は 0.01, 0.05 以下であり、有意な結果である可能性が高く、形容詞とフォロワー数との間に一定の関連性があると考えられる。

続いて Table 3 の順位相関係数の下位 10 件の結果について触れる。下位 10 件の順位相関係数は全て負であり、フォロワー数との負の相関が表れている。全体の傾向を見ると、上位 10 件と比較してやや解釈が難しい形容詞が多く含まれている。例えば、最も負の相関が強い「久しい」は、対象インフルエンサーのコンテンツに久しく触れていない状態を間接的に示しており、これがフォロワー数順位に負の影響を及ぼす、ということを示している。関連して、「寂しい」の否定語（寂しくない）は、対象インフルエンサーの休暇・活動休止等に際して、Tweet 投稿者が自分の寂しさを否定する形で気持ちを表現する文脈で多々用いられていた。このように下位 10 件は、対象インフルエンサーに対する直接的なネガティブ表現は限られているものの、間接的な形でフォロワー数への負の影響が示唆される評価・感情表現を示す形容詞が出現している。また、一見するとポジティブな表現に見られる「暖かい」（人柄等に対する「温かい」という表現の誤用）「偉い」「大人っぽい」といった形容詞も負の相関を持つことを示している。

以上で述べたように、直感的には判断しづらい世間からの評価・感情表現を形容詞という形で集約し、人気との関連性を具体的に表現することができている。

相関係数の絶対値は大きいもので 0.4 程度であるものの、特定の形容詞が有意に正・負の相関を持つことが分かった。

## 5.2. 感情とフォロワー数の順位相関係数

Table 4 は感情の出現割合とフォロワー数の順位相関係数を表している。

各感情について、5.1 節で述べた形容詞の順位相関係数と比べて絶対値が小さく、形容詞と比較すると人気に対する相関が弱いということが分かる。その中でも比較的有意に相関が大きかったのが「恥 (Shame)」の感情である。これは、恥じらいの感情が含まれる語が多く出現するほど、フォロワー数が多いということを示している。感情分析の際に「恥 (Shame)」に分類された形容詞の例としては、「恥ずかしい」や「照れくさい」などがあり、それらの形容詞が含まれている Tweet の原文を調べると、対象インフルエンサーが何かしらに失敗して恥ずかしがる様子を揶揄したり、恥ずかしがる様子に可愛げを覚えたり、といったようなポジティブな内容の意見が多く含まれていた。

5 章をまとめると、インフルエンサーを対象とした分析においては、対象に向けられた評価・感情表現を表す特定の形容詞が Twitter フォロワー数に対して相関を持っていることが分かった。また、感情分析によって抽出した感情とフォロワー数との間の相関は、形容詞と比べると弱いものの、「恥 (Shame)」のように一定影響を与えうる感情も確認された。

## 6. 考察

5 章の実験結果について、研究目的である人気の要因説明に踏み込んで考察をする。Twitter 上の対象インフルエンサーと共に起る形容詞について、フォロワー数との順位との間に有意な相関が見られ、人気に対して一定の影響を持つ要素として解釈が可能と考える。

例えば、5.1 節では食に関する感想が多く含まれており、食という人類共通的に興味関心を持ちやすい分野についてコンテンツを配信すると、その感想がポジティブ (美味しい、旨い)・ネガティブ (不味い) のどちらかに関わらず、人気が高まりやすいということを示唆している。また、インフルエンサーのパフォーマンス (例えば、即興での歌唱、突出した運動能力、異性を惹きつける行動・発言など) に対する賛辞・尊敬・驚嘆を示す評価 (凄い、やばい) を集めているインフルエンサーも、より人気が高まりやすいということも示している。また、これらの示唆が YouTuber の人気獲得のための戦略策定に役立てられる可能性がある。例えば、今後作成する動画コンテンツの中に食に関する内容を多く取り入れたり、競合に対して比較優位性のある分野 (例えば、歌唱、運動、トークなど) のコン

텐츠作成に注力したりして、大衆から「凄い」「やばい」といった賛辞・尊敬・驚嘆を示す評価を集めるようにすることなどが考えられる。また、「久しい」「寂しくない」といった、YouTuber のフォロワーとの関わりが減少することにより発生しがちな印象・感想を抑制する施策を行うことによって、フォロワー数の維持に繋げられる可能性がある。

5.2 節の感情に関する分析では、「恥 (Shame)」という感情に紐づく様子、具体事例を挙げると対象インフルエンサーの恥ずかしがる様子等が多く見られる方が大衆からの人気を集めやすいということを示している。これも YouTuber の人気獲得のための戦略形成に役立てられる可能性がある。例えば、キャラクターとして隙があるような部分を自己開示するように発信活動をするなど、大衆から「恥 (Shame)」を含む印象を強める活動を広げることでフォロワー数の増加を狙える可能性がある。

より詳細な分析、戦略策定を進めるためのアプローチとして、トピックモデルを併用し、相関が強い各形容詞・感情がどのようなトピックが紐づいているかを調べるなどが考えられる。

## 7. まとめと今後の展望

本研究では汎用性・解釈性の高い人気の要因分析手法を提案することを目的とし、Twitter データを用いて日本のインフルエンサーに対する大衆の意見を抽出し、人気の指標となる Twitter 上のフォロワー数との関係性を調べた。具体的な分析手法としては、意見に関する重要な指標である形容詞と、それに紐づく感情を Tweet から分析要素として抽出し、各要素の出現頻度の順位と、フォロワー数の順位の関係性を、Spearman の順位相関係数を算出することで分析した。

結果として、一部の形容詞・感情について相関係数 0.3-0.4 程度でフォロワー数順位と有意に相関していることが分かり、YouTuber の人気に対して一定の説明力を持つ指標であることが示唆された。

今後の展望としては、今回得られたような人気との相関を強く示す感情について、紐づく元意見の内容をさらに詳細に分析することで、対象インフルエンサーの人気を向上させるための具体的なアクションプランに結び付くような、さらなるインサイトを発見する手法について検討したい。また、本研究では人気を説明するための要素として形容詞と感情を扱ったが、本研究の提案手法であれば他の形式のデータであっても容易に導入可能であるため、人気にクリティカルに影響する別の要素がないか適用範囲を広げることも検討したい。また、提案手法の対象範囲はインフルエンサーだけでなく、ヒト (例: 政治家、スポーツ選手)・モノ

(例：ブランド，生活消耗品，サービス)・コト(例：イベント)など他のセグメントにも応用可能なので，人気獲得の需要が大きいフィールドにも分析範囲を広げたい。

ただし，今回の手法には課題も残っている．例えば，形容詞の抽出条件として，対象インフルエンサーと同じ文中で共起するものを抽出したが，この手法だと対象以外に向けられている無関係な形容詞が混じってしまう．ノイズを低減するための手法として係り受け解析を利用することも可能ではあるが，本研究では分析用のデータ数の確保等の観点から利用を見送った．しかし，今後より精度の高い結果を得るため，係り受け解析の利用も視野に含め，ノイズ低減のために手法を改善する必要がある．また，感情分析には ML-Ask を用いたが，このモデルはキーワードベースで感情を分類しており，データベースの語数が 2100 語と限られていたため，感情分類できた形容詞が限定的であった．今後，より精度の高い感情分類ができるように，別モデルや感情表現辞典をベースとした手法を検討したい。

## 謝辞

本研究の一部は JST・AIP (JPMJCR22U4)の支援を受けた。

## 参考文献

- [1] Davenport, T. H., & Beck, J. C., "The attention economy," *Understanding the New Currency of Business*. Harvard Business Press, 2001.
- [2] Cambria, E., Schuller, B., Xia, Y., & Havasi, C., "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent systems*, vol. 28, no. 2, pp. 15-21, 2013.
- [3] Pak, A., & Paroubek, P., "Twitter as a corpus for sentiment analysis and opinion mining," *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 2010.
- [4] Ravi, K., & Ravi, V., "A survey on opinion mining and sentiment analysis: tasks, approaches and applications," *Knowledge-based systems*, vol. 89, pp. 14-46, 2015.
- [5] Ren, J., Shen, J., & Kauffman, R. J., "What Makes a Music Track Popular in Online Social Networks?," *Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion)*, pp. 95-96, 2016.
- [6] Khosla, A., Das Sarma, A., & Hamid, R., "What makes an image popular?," *Proceedings of the 23rd international conference on World wide web (WWW'14)*, 2014.
- [7] Spearman, C., "The proof and measurement of association between two things," *The American Journal of Psychology*, vol. 15, no. 1, p. 72-101, 1904.
- [8] Yan, H., Dai, J., Ji, T., Qiu, X., & Zhang, Z., "A Unified Generative Framework for Aspect-based Sentiment Analysis," *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (IJCNLP'21)*, vol. 1 (Long Papers), pp. 2416-2429, 2021.
- [9] Hazarika, D., Zimmermann, R., & Poria, S., "Misa: Modality-invariant and-specific representations for multimodal sentiment analysis," *Proceedings of the 28th ACM international conference on multimedia (MM'20)*, pp. 1122-1131, 2020.
- [10] Cao, Q., Shen, H., Gao, J., Wei, B., & Cheng, X., "Popularity prediction on social platforms with coupled graph neural networks," *Proceedings of the 13th International Conference on Web Search and Data Mining (WSDM'20)*, pp. 70-78, 2020.
- [11] Liao, D., Xu, J., Li, G., Huang, W., Liu, W., & Li, J., "Popularity prediction on online articles with deep fusion of temporal process and content features," *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'19)*, vol. 33, no. 1, pp. 200-207, 2019.
- [12] Yamasaki, T., Sano, S., & Aizawa, K., "Social Popularity Score: Predicting Numbers of Views, Comments, and Favorites of Social Photos Using Annotations," *Proceedings of the First International Workshop on Internet-Scale Multimedia Management (WISMM'14) in conjunction with ACM Multimedia*, pp. 3-8, 2014.
- [13] Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V., "Support vector regression machines," *Advances in neural information processing systems*, vol. 9, p. 155-161, 1996.
- [14] Kudo, T., "Mecab: Yet another part-of-speech and morphological analyzer," 2005. [Online]. Available: <http://mecab.sourceforge.net/>. [Accessed 10 1 2022].
- [15] "mecab-ipadic-NEologd," [Online]. Available: <https://github.com/neologd/>. [Accessed 3 1 2023].
- [16] 岩成 達哉, 吉永 直樹, 豊田 正史, 喜連川 優, "多様な手がかりを用いた形容詞に基づく概念語の順序付け," 第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM'16), pp. 2-3, 2016.
- [17] Ptaszynski, M., Dybala, P., Rzepka, R., Araki, K., & Masui, F., "ML-Ask: Open source affect analysis software for textual input in Japanese," *Journal of Open Research Software*, vol. 5, no. 1, p. 16, 2017.
- [18] 中村 明, "感情表現辞典 東京堂," 1996.
- [19] Paul, N., & Sui, M., "I Can Feel What You Feel: Emotion Exchanges in Twitter Conversations between Candidates and the Public," *Journal of Political Marketing*, vol. 21.2, pp. 175-195, 2022.

付録 A 実験対象の YouTuber 一覧

順位	対象	フォロワー数	収集Tweet数	備考
1	HIKAKIN	6,240,545	190,034	
2	はじめしゃちょー	4,814,052	33,049	
3	てつや(東海オンエア)	2,194,181	25,939	東海オンエア所属
4	ブライアン	2,152,657	3,075	
5	キヨ	1,810,183	24,185	ゲーム実況者
6	シルクロード(フィッシャーズ)	1,672,146	11,590	フィッシャーズ所属
7	りょう(東海オンエア)	1,592,112	12,037	東海オンエア所属
8	としみつ(東海オンエア)	1,564,353	2,238	東海オンエア所属
9	虫眼鏡(東海オンエア)	1,440,847	5,230	東海オンエア所属
10	DJ社長(Repezen Foxx)	1,425,067	59,270	Repezen Foxx所属
11	ンダホ(フィッシャーズ)	1,404,828	16,848	フィッシャーズ所属
12	からめる	1,393,883	26,102	漫画家、動画クリエイター
13	ヒカル	1,383,838	38,528	
14	しばゆー(東海オンエア)	1,363,841	18,283	東海オンエア所属
15	フィッシャーズ	1,327,246	10,788	フィッシャーズ公式アカウント
16	白上フブキ	1,283,426	24,134	バーチャルYouTuber事務所「ホロライブ」所属
17	湊あくあ	1,275,804	36,347	バーチャルYouTuber事務所「ホロライブ」所属
18	テオ(スカイピース)	1,255,648	52,427	スカイピース所属
19	カンタ	1,254,926	42,763	水溜りボンド
20	SEIKIN	1,249,383	17,092	
21	兔田ぺこら	1,236,463	51,670	バーチャルYouTuber事務所「ホロライブ」所属
22	マサイ(フィッシャーズ)	1,218,350	18,855	フィッシャーズ所属
23	☆イニ☆(じん) (スカイピース)	1,169,916	29,981	スカイピース所属
24	フワちゃん	1,079,737	43,692	
25	葛葉	1,078,356	249,193	Vtuber
26	あやなん	1,073,605	11,978	
27	HAACHAMA	1,052,097	44,949	ホロライブ
28	トミー	1,039,076	5,283	水溜りボンド
29	ゆめまる	1,031,550	26,735	
30	にじさんじ	1,021,124	281,345	バーチャルライバーグループ「にじさんじ」公式アカウント
31	ジェル(すとぷり)	1,014,878	69,210	エンターテインメントグループ「すとぷり」所属
32	すしらーめんりく	913,367	27,371	
33	叶(にじさんじ)	910,418	66,118	バーチャルライバーグループ「にじさんじ」所属
34	赤髪のともし	888,563	23,305	ゲーム実況者
35	スカイピース	858,003	40,924	スカイピース公式アカウント
36	弟者	853,050	83,915	ゲーム実況者
37	莉犬くん(すとぷり)	831,581	92,411	エンターテインメントグループ「すとぷり」所属
38	DJふおい(Repezen Foxx)	782,657	30,821	Repezen Foxx所属
39	さとみ(すとぷり)	764,163	16,216	エンターテインメントグループ「すとぷり」所属
40	桐崎栄二	757,483	2,662	
41	ころん(すとぷり)	747,082	29,157	エンターテインメントグループ「すとぷり」所属
42	ななもり。(すとぷり)	730,954	38,081	エンターテインメントグループ「すとぷり」所属
43	剣持刀也(にじさんじ)	721,500	16,868	バーチャルライバーグループ「にじさんじ」所属
44	加藤純一	717,064	30,205	
45	ポッキー	714,179	42,982	ゲーム実況者
46	DJ銀太(Repezen Foxx)	708,700	21,425	Repezen Foxx所属
47	P丸様。	706,679	52,438	
48	ときのそら	703,163	13,943	バーチャルYouTuber事務所「ホロライブ」所属
49	癒月ちよこ	701,938	71,105	バーチャルYouTuber事務所「ホロライブ」所属
50	ロボ子さん	668,286	53,841	バーチャルYouTuber事務所「ホロライブ」所属
51	月ノ美兔	645,585	40,041	バーチャルライバー