

# 人気アカウント分析 ～Twitter において人気であるためには～

石川県立金沢泉丘高等学校 2年

大野 諒英 柿本 理一 小林 賢太郎 新保 幸輝 丹羽 剣竜 物部 優太

## 1. 要旨、概要

### i. 要旨

近年、情報化社会となり情報の価値が高まりつつあるので、私たちは情報伝達手段として Twitter に注目し、アカウントを分析した。私たちは各アカウントの「いいね」数の中央値を人気度と定義し、集めたデータをもとにさまざまな指標との相関を調べた。その結果、次のような知見が得られた。

- (1) 社会的地位が高いアカウントの人気度は高い
- (2) 人気度が高くなる理想のツイート頻度が存在する
- (3) トレンドワードを含まないツイートは人気度が高い

ということがわかった。この結果を、実際にアカウントを作成し検証したい。

### ii. 背景

日本でも、社会の情報化に伴い Twitter<sup>i</sup>をはじめ SNS が広く普及している。SNS に関する研究は多数あり、様々な面で応用できることが知られている。先行研究では使用言語における Twitter の利用方法や特徴の差を調べられていたが、私たちは Twitter でより多くの人に情報を発信する方法を探りたいと考え、Twitter での人気アカウントの分析が有効なのではないかと考えた。なぜなら人気アカウントはその人気さ故に多くの人に自らの情報を発信することが可能となる。そこで私たちは人気のあるアカウントにおいて、その傾向、共通点を多角的に調べることにした。

私たちが研究に Twitter を選んだ理由は3つある。

1つ目は多くの情報がリアルタイムで更新されることである。常に私たちは新しい情報を手に入れることができるため、情報を集めることが目的であるアクティブユーザーも数多く存在すると考えた。

2つ目は簡単に情報を世界に発信することが可能であることだ。Twitter ではつぶやき1つで些細な情報から重大な情報まで簡単に情報を発信することができる。

そして3つ目は日本で若者の間での利用率が高いことだ。私たちと同じ20代以下での SNS 利用率において Twitter が高い割合を占めていた。(図1)

## 2. 研究目的

この研究の目的はツイート<sup>ii</sup>が多くの人の目にとまるような人気アカウントを作成することである。Twitter をはじめ SNS では、不特定多数の人に情報を伝えることが可能であり、それらの情報は受け取った人々に影響を与え、その考えや心を動かすことで、行動を変えることさえも可能となる。人気アカウントはより多くの人に情報を伝えることができるので、より多くの人に影響を与えられる。そのようなアカウントを作成し、操作することができれば日本だけでなく世界中へ影響力を持つアカウントを作成することができるはずだ。

そのような人気アカウントには、ツイートやプロフィールの中に多くの共通点があると考えられるので、複数の人気アカウントが共通して持つ特徴や傾向を見つけることが人気アカウントの作成につながると考えられる。まず、人気度というものを定義し、複数のアカウントについて、人気度と様々な要素を比較し人気アカウントの傾向や共通する特徴を調べる。私たちがそれらの結果を踏まえることで、研究の結果に

基づいた人気アカウントを作成することができるはずだ。また、その人気性を利用してネット広告などの商業利用や、他の分野への応用を考えることも可能となるだろう。

私たちは、人気度と比較する要素の中でも特に、ツイートの頻度、ツイート内の最新情報の有無、ユーザーの社会的地位の三点がアカウントの人気度に関係があるという仮説を立てた。

それぞれの仮説を定めた理由は次のとおりである。まず、ツイートの頻度について、ツイートの頻度が低くなると常に更新されるたくさんの人たちのツイートの中で自分のツイートが目にとまる可能性は低くなっていく。それに伴って、他人が進んでフォロー<sup>iii</sup>する可能性も低くなるだろう。逆に頻度が高ければ、自分のツイートがより人目に触れるのでフォローしてくれる人が増える可能性は高くなる。また、フォロワーが多いと、それだけ多くの人に見てもらえることになるので私たちの目指す人気アカウントとなり得る。

次に、ツイート内の最新情報の有無について、先ほども述べたように Twitter では多くの情報が常に更新されている。最新情報は特にすぐに情報を手に入れたらというユーザーの関心を惹くものであり、少なからずツイートを見る人は増えると考えられるだろう。

そして、ユーザーの社会的地位について、社会的地位の高い人は多くの人に知られている可能性が高く、その場合には一般的なユーザーに比べて進んでフォローされると考えられる。芸能人を例として挙げると、テレビに多く出演している芸能人は Twitter のアカウントとは別に高い知名度を持っているため、同様に Twitter でのフォロワー<sup>iv</sup>数も増やすことができると考えられる。

### 3. 研究方法

私たちの研究では、信憑性及び一般性を持たせるために多くのデータを取り扱う必要がある。それにあたって、私たちは短時間で多量のデータを取得する方法を考えなければならなかった。

そこで私たちは、「Python」というプログラミング言語を使うことにした。この言語は初心者でも比較的使いやすく、また、大量のデータを

扱う際によく用いられる。また、私たちの中に研究以前から Python を習得している人がいたということもこの言語を選んだ理由の一つに挙げられる。私たちはこの Python を用いて「Twitter API」<sup>v</sup>にアクセスし、情報を入手する。私たちは研究に先駆け、API 操作のためのアカウントを作成した。

実際に研究を始めるにあたり、まず人気アカウントというものを定義する必要がある。そこで、私たちはまずこの研究における人気の定義について考えた結果、「いいね」<sup>vi</sup>数の中央値を人気度と定義した。なぜなら、「いいね」とは他人から受ける評価の指数であり、リツイート<sup>vii</sup>よりも好意的な意味を含む場合が多いと考えたからである。また中央値は突発的な値を持つような外れ値の影響を受けにくく、恒常的に人気なアカウントを調査できると考えたからである。

実際にデータは Twitter API から、users/search APIにおいて検索キーワードとして日本語の助詞を用いることにより、無作為にユーザーを抽出した。そのユーザーのデータを statuses/user\_timeline API に指定することで最大 200 件の各ユーザーのツイートを抽出した。また、後述するトレンドワードを含むツイートに関してはタイムライン上に表示されるトレンドワードを search/tweets API に対して入力することによりツイートの情報を入手した。集計期間と件数について、2016 年 12 月 12 日時点のユーザーを 2399 件、ツイートに関しては 2016 年 12 月 12 日付近のツイートを 335215 件取得した。

分析では調査した人気度と 6 つの指標についてそれぞれ比較した。6 つの指標とは、ツイート頻度、リツイート数、フォロワー数、リプライ<sup>viii</sup>率、公式アカウント<sup>ix</sup>か否か、トレンドワードを含むツイートをするか否か、である。

この分析におけるツイート頻度とは、最新百件のツイートが何日間によって行われているかを示したものである。つまり、一日当たり何件ツイートされているかということである。この値が大きければ、ツイート頻度が高いといえる。

リプライ率とは、ツイート総数に対するリプライ数の割合である。

トレンドワードとは、データ集積当時のツイートに多く現れた言葉であり、これらは

Twitter 社から提供されている。トレンドワードとなる言葉は Twitter の内外で話題となった言葉であることが多く、これらが最新情報に結びついているのではないかと考え、人気度と比較した。

また、分析においては獲得した集団をそれぞれの指標に即して順に並べ、第 1 四分位数、中央値、第 3 四分位数によって 4 つの集団に分けて分析を行った。

#### 4. 結果

##### i. ツイート頻度

グラフ 1 より、4 つの集団のうち、ツイート頻度の最も高い集団、最も低い集団の人気度が低いことが分かる。

##### ii. リツイート数

グラフ 2 より、リツイート数が多い集団は少ない集団に比べ、人気度の平均が高い。

##### iii. フォロワー数

グラフ 3 より、フォロワー数が多い集団は少ない集団に比べ、人気度の平均が高い。

##### iv. リプライ率

グラフ 4 より、リプライ率の高い集団は低い集団に比べ、人気度の平均が高い。

##### v. 公式アカウントか否か

公式アカウントの人気度の平均は 707、非公式アカウントの人気度の平均は 180 であった。

##### vi. トレンドワードを含むツイートをするか否か

仮説に反し、トレンドワードを含むツイートを行う集団に比べ、そのツイートを行わないアカウントの集団が人気度の平均値が高かった。このことから、トレンドワードは、頻繁にツイートされている言葉というだけであり、ツイートを見る人の心を動かすと

いう意味とは一致しないからだと考えられる。

#### 5. 考察

##### i. ツイート頻度

結果から、4 つの集団のうち、ツイート頻度の最も高い集団、最も低い集団の人気度が低いことが分かる。このことから、理想のツイート頻度があると考えられる。この理由としては、頻度が低すぎるとはツイートが閲覧されにくく、高すぎるとは閲覧する人が煩わしいと感じるためと思われる。よって、よりツイート頻度の区分を細かくすることにより、理想のツイート頻度の値を得られると考えられる。

##### ii. リツイート数

結果より、リツイート数の増加は人気度の増加に影響するといえる。この理由としてはリツイートと「いいね」は同時に行われやすいからと考えられる。

##### iii. フォロワー数

結果より、フォロワー数の増加は人気度の増加に影響するといえる。これは、フォロワーが増えるとツイートの閲覧者が増えるため、「いいね」がされやすくなり、人気度は高くなると考えられる。

##### iv. リプライ率

結果から、リプライ率が増加することで人気度が減少する、つまり、リプライ率の減少が人気度の増加に影響するといえる。これは、リプライは特定のアカウントに向けられたツイートであるため、通常のツイートよりも「いいね」がされにくく、よってリプライ率の低いアカウントは不特定多数に向けられたツイートが多いため、人気度が高いと考えられる。

##### v. 公式アカウントか否か

社会的地位が高いと考えられる公式

アカウントはそうでないアカウントと比べ、人気度が高いといえる。これは、そもそも公式アカウント自体は著名であるということがツイッター社に認証されたアカウントであるため、一般の人々からの知名度が高く、それに加えツイート内容がメディア等に取り上げられる場合がしばしばあるため、人気度は高くなると考えられる。

vi. **トレンドワードを含むツイートをするか否か**

結果から、トレンドワードは、頻繁にツイートされている言葉というだけであり、ツイートを見る人の心を動かすとは限らないからだと考えられる。また、同じ内容のツイートが増えるため、見る人がその内容に飽きてしまうため、またトレンドワードを含むツイートに「いいね」がされるとしてもツイートを行ったアカウントの数が多いため「いいね」が分散されるため、という他の要因も大いに影響していると考えられる。

**6. 結論**

私たちは、ツイートの頻度、ツイート内の最新情報の有無、ユーザーの社会的地位の三点がアカウントの人気度に関係があるという仮説を立てたが、分析結果より社会的地位の高いユーザーは人気度が高いということが分かった。また、ツイートの頻度には、高すぎず、低すぎない、理想の頻度が存在するとわかった。さらに、私たちの仮説に反して、トレンドワードを含まないツイートの方がそれを含むツイートより人気度が高かった。

私たちは、この結果から人気度の高いアカウントの特徴をつかむことができた。私たちは、多くの人に情報を発信するために、ツイッターのアカウント分析を行ったが、その検証には至らなかった。

今後は、検証用のアカウントを作成して、この分析結果の通りにアカウントを操作すると、人気度の高い、多くの人に情報を発信できるアカウントが作成できるか検証したい。

**7. 参考文献**

竹村光, 田島隆史, 情報発信の対象範囲に基づく Twitter ユーザーの分類, DEIM Forum 2013B1-6, [https://www.jstage.jst.go.jp/article/tjsai/30/1/30\\_30\\_393/\\_pdf](https://www.jstage.jst.go.jp/article/tjsai/30/1/30_30_393/_pdf), (参照 2016-05-02)

池田和史, 服部元, 松本一則, 小野智弘, 東野輝夫, マーケット分析のための Twitter 投稿者プロフィール推定方法, 情報処理学会論文誌 コンシューマ・デバイス&システム, 2012, Vol. 2, No. 1, 82-93,

[https://ipsj.ixsq.nii.ac.jp/ej/index.php?action=pages\\_view\\_main&active\\_action=repository\\_action\\_common\\_download&item\\_id=79856&item\\_no=1&attribute\\_id=1&](https://ipsj.ixsq.nii.ac.jp/ej/index.php?action=pages_view_main&active_action=repository_action_common_download&item_id=79856&item_no=1&attribute_id=1&) (参照 2016-05-02)

石井健一, マイクロブログ Twitter における日本人利用者の特徴, Department of Social Systems and Management

Discussion Paper Series, 2011, no. 1277,

<https://www.sk.tsukuba.ac.jp/SSM/libraries/pdf/1276/1277.pdf> (参照 2016-10-24)

総務省情報通信国際戦略局情報通信政策課情報通信経済室. ”社会課題解決のための新たな ICT サービス・技術への人々の意識に関する調査研究—報告書—”. 2015-03

[http://www.soumu.go.jp/johotsusintokei/linkdata/h27\\_06\\_houkoku.pdf](http://www.soumu.go.jp/johotsusintokei/linkdata/h27_06_houkoku.pdf)

(参照 2016-04-21)

**8. 謝辞**

今回の研究を行うにあたって、多くの先生方や大学の教授の方、また留学生の方々からアドバイスや指導をしていただきました。ご協力ありがとうございました。

## 1. 図表・グラフ

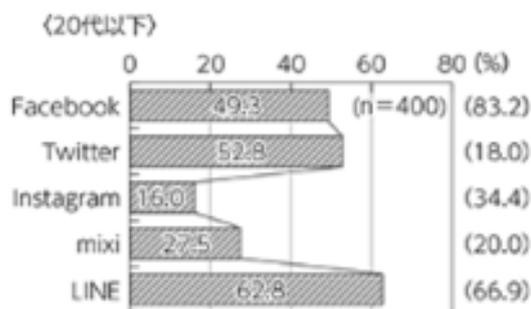
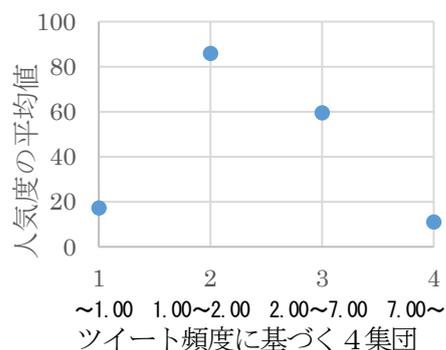


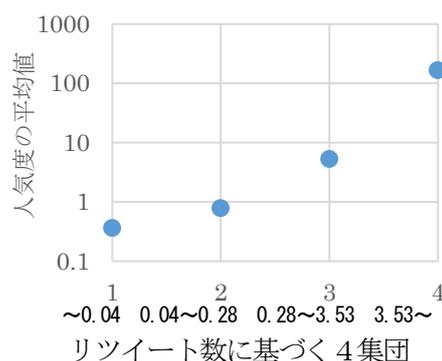
図1 総務省ホームページより

[http://www.soumu.go.jp/johotsusintokei/linldata/h27\\_06\\_houkoku.pdf](http://www.soumu.go.jp/johotsusintokei/linldata/h27_06_houkoku.pdf)

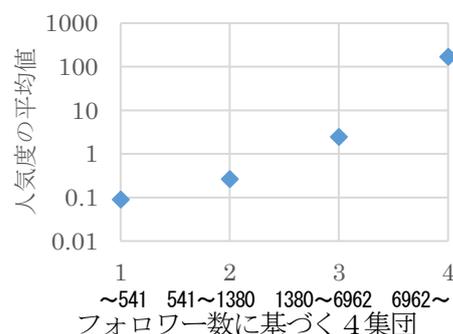
グラフの横軸は、それぞれの項目において、数値順に並べ、第1四分位数・中央値・第3四分位数で区切り、4つの集団に分け、数値の低い順に1・2・3・4とした。



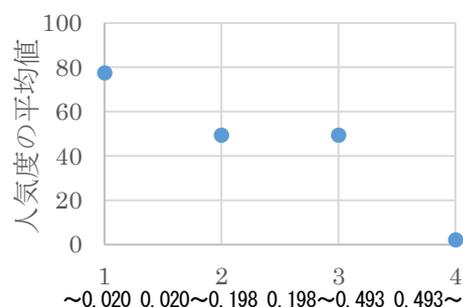
グラフ1 ツイート頻度と人気度の比較



グラフ2 リツイート数と人気度の比較



グラフ3 フォロワー数と人気度の比較



グラフ4 リプライ率と人気度の比較

- i ユーザーが「つぶやき」と呼ばれる140字以内の短い記事を書き込み、ほかのユーザーがそれを読み、返信をすることでコミュニケーションが生まれるインターネット上のサービス。
- ii 140文字以内に限定された情報を発信すること。また、その情報。
- iii 他人の投稿（ツイート）を自分のページで見られるよう登録すること。
- iv フォロワーした人。自分の投稿（ツイート）を見られるよう登録した人。

- v 自分のアカウント情報を提示することで、プログラムによってTwitter内の情報を入手することのできるサービス。
- vi 他のユーザーに対して好意を示す働きかけ。
- vii 他のユーザーのツイートを自分のアカウントから再投稿すること。
- viii ツイートに対して行う返信。
- ix Twitter社から本人と認証されているアカウント。