

# テレビアニメ作品に関する SNS 上の情報拡散傾向と感情の関係

石倉直樹<sup>1</sup> 土屋雅稔<sup>1</sup> 吉田光男<sup>2</sup>

<sup>1</sup> 豊橋技術科学大学大学院 情報・知能工学専攻

<sup>2</sup> 筑波大学 ビジネスサイエンス系

{ishikura,tsuchiya}@is.cs.tut.ac.jp, mitsuo@gssm.otsuka.tsukuba.ac.jp

## 概要

Twitter (現在は X) 上でのテレビアニメ作品の情報拡散傾向は多様であり, その作品に関するツイートを行うユーザが増加する作品もあれば, 減少していく作品もある. このように情報拡散傾向が異なる場合, ユーザの行動も異なることが予想される. そのため, 本稿ではテレビアニメ作品を対象に情報拡散傾向の違いによる, ツイートに含まれる感情の違いに注目して分析を行う. 最初に情報拡散傾向の種類を週間ツイートユーザ数の推移データによるクラスタリングと, 平均週間ツイートユーザ数によるクラスタリングにより定義した. そして, 感情分析モデルを用いて, 情報拡散傾向の種類ごとにツイート内容に対して感情分析を行なった. その結果, 週間ツイートユーザ数の推移傾向が上昇傾向にある場合, 他の傾向よりもネガティブツイートの割合が高いなど, 情報拡散傾向の違いにより, ツイートに含まれている感情にも違いが現れることを確認した.

## 1 はじめに

近年, 動画配信サイトの普及により, テレビアニメ作品をリアルタイムではなく, 放送終了後も視聴することが可能となった. また, Peihua らの研究によると, SNS 上でテレビ番組の口コミが拡散されるほど, オンデマンドでの視聴が増加することがわかっている [1]. そのため, SNS 上での口コミの影響によりテレビアニメ作品を動画配信サイトで視聴し, その後リアルタイムの視聴に参加するユーザが存在すると考えられる. 実際に, 図 1 のテレビアニメ作品 A の放送期間中における週ごとのツイートしたユーザ数を見ると, ユーザ数が増加傾向にあることがわかる. これより, 放送期間の途中からテレビアニメ作品の視聴, および情報拡散に参加したユーザが多く存在すると考えられる. 逆にテレビアニメ作品 B のように, 週ごとにツイートしたユーザが減少している作品も

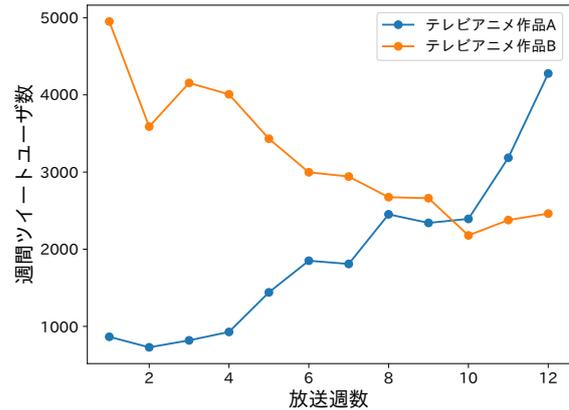


図 1 2つのテレビアニメ作品の週ごとのツイートユーザ数

ある. このようにテレビアニメ作品によって情報拡散傾向の違いがあり, それによって SNS 上でのユーザの行動にも違いが現れることが予想される. 例えば, 図 2 と図 3 はそれぞれテレビアニメ作品 A と B の週ごとの感情別ツイート数であり, 両作品を比較すると上昇傾向にある図 2 の作品の方が, 下降傾向にある図 3 の作品よりもネガティブなツイートが多いことがわかる.

また, テレビアニメ作品は多くの作品が 3 ヶ月や半年といった期間で毎週放送されるため, 定期的にイベントが発生する場合の情報拡散傾向を複数観察することができ, それらの違いを比較しやすい. さらに, Twitter (現在は X) には, テレビアニメ作品を好むユーザが多く存在することが榊らの研究 [2] で示されており, 分析対象として適している.

そのため, 本稿では, テレビアニメ作品を対象に, Twitter 上での情報拡散傾向の違いによるツイートに含まれている感情の違いを分析する. その比較により, 情報拡散傾向の違いによるユーザ行動の違いを明らかにする.

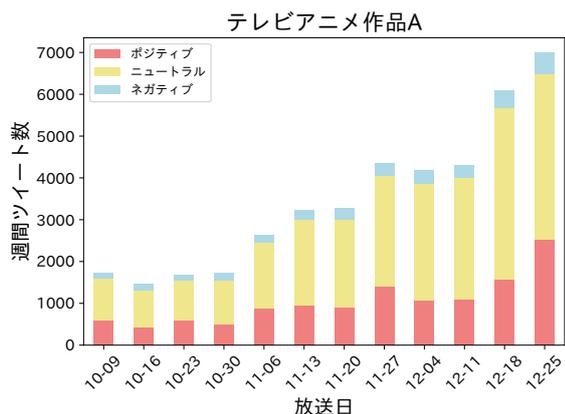


図2 テレビアニメ作品Aの週ごとの感情別ツイート数

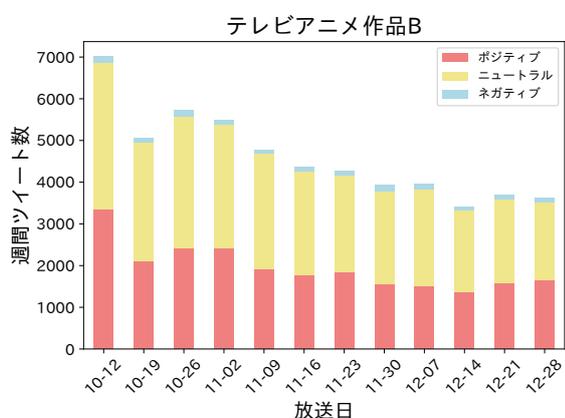


図3 テレビアニメ作品Bの週ごとの感情別ツイート数

## 2 関連研究

Chawlaらは、SNS上では、joyやcontempt, guilt, distressといった感情を含む投稿が拡散されやすいことが報告している[3]。相川らは、SNS上でのテレビ番組の情報拡散と投稿に含まれる感情の関係を分析した研究を行い、深夜アニメ番組に関するツイートに着目してツイートを分析し、前週の実況ツイート内のマイナス表現と、次週の視聴率には正の相関があることを示している[4]。これらの研究では、投稿に含まれる感情と情報拡散の関係を分析しており、これらの結果より、情報拡散傾向の違いによってツイートに含まれる感情にも違いが現れると予想できる。そのため、本稿ではその点に注目して分析を行う。

## 3 データの収集

本稿での分析には、多くのテレビアニメ作品とTwitterのデータを必要とする。そのため、Wikipediaの「日本のテレビアニメ作品一覧(20年代前

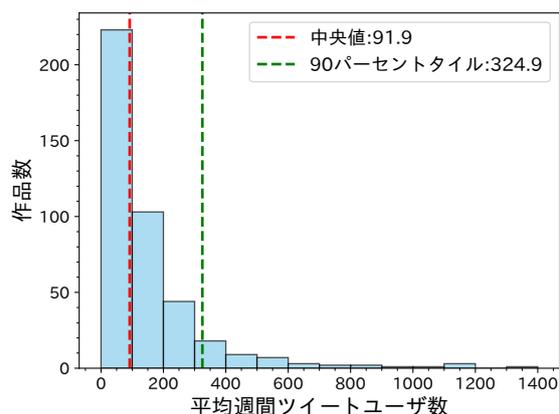


図4 放送期間が11~13週のテレビアニメ作品の平均週間ツイートユーザ数の分布

半)」<sup>1)</sup>から、2020年1月1日から2023年4月1日までに放送開始し終了した652作品のテレビアニメ作品データを取得した。そして、本研究では情報拡散に重点を置くため、Twitter REST API v1.1によって収集された日本語のリツイートデータから、分析に使用するデータの抽出を行った。テレビアニメ作品に関するリツイートデータの抽出方法は、各作品の公式アカウントのツイートに1番多く含まれていたハッシュタグを、各作品の代表的なハッシュタグとし、それが含まれているリツイートデータを抽出した。

その結果、APIによるデータの収集ができなかった期間を除く、579作品のテレビアニメ作品のリツイートデータを取得した。抽出したリツイートデータの合計は384,629,894件で、1作品あたり663,154件、抽出したオリジナルツイートデータの合計は12,562,702件で、1作品あたり21,659件である。

## 4 情報拡散傾向のクラスタリング

情報拡散傾向として上昇傾向や下降傾向などが考えられるが、単純なルールや条件によってそれらに分類することは難しい。そのため、データの前処理後、2つの観点からクラスタリングを行い、情報拡散傾向の種類を定義する。その2つの観点とは、週に特定のテレビアニメ作品についてツイートしたユーザ数(以後、週間ツイートユーザ数)の推移データと、平均週間ツイートユーザ数である。

1) [https://ja.wikipedia.org/wiki/日本のテレビアニメ作品一覧\\_\(20年代\\_前半\)](https://ja.wikipedia.org/wiki/日本のテレビアニメ作品一覧_(20年代_前半))

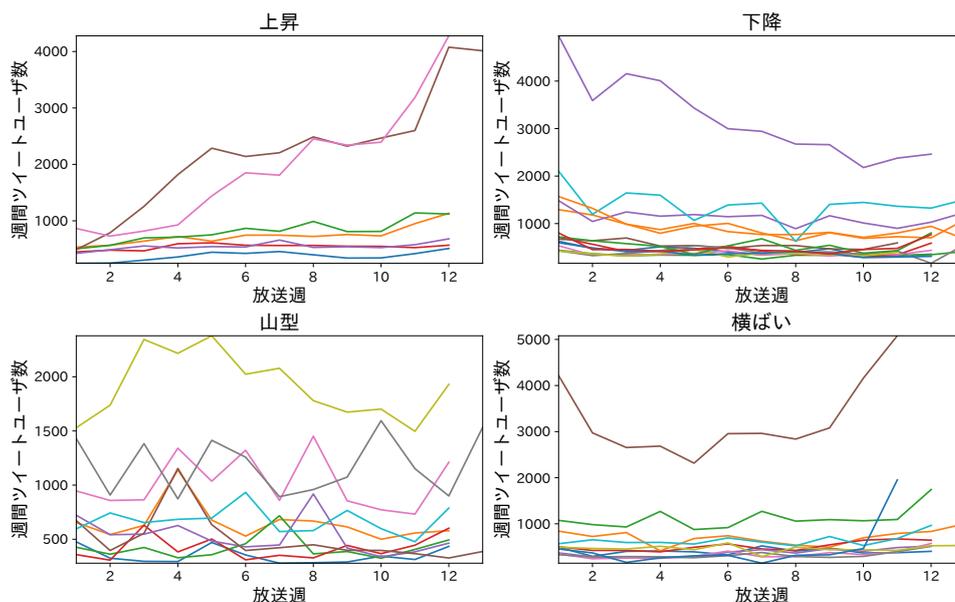


図5 週間ツイートユーザ数推移データのクラスタリング結果

#### 4.1 データの前処理

まず、テレビアニメ作品の放送期間にはばらつきがあるため、放送期間が11~13週のテレビアニメ作品を対象としてクラスタリングを行う。

また、テレビアニメ作品ごとの平均週間ツイートユーザ数には大きな差がある。図4のテレビアニメ作品の平均週間ツイートユーザ数の分布(1500以上は除去してプロットしている)を見ると、分布がべき乗則のようになっており、50%以下の作品は平均週間ツイートユーザ数が91.9で、90%以下の作品は324.9であることがわかる。このように作品ごとの平均週間ツイートユーザ数には大きな差があるため、平均週間ツイートユーザ数が上位10%のテレビアニメ作品を対象としてクラスタリングを行う。

つまり、放送期間が11~13週のテレビアニメ作品を抽出した後、さらに平均週間ツイートユーザ数が上位10%のテレビアニメ作品を抽出した計43作品を対象にクラスタリングを行う。

#### 4.2 週間ツイートユーザ数推移データのクラスタリング

週間ツイートユーザ数推移データをクラスタリングし、推移傾向を分類する。まず、テレビアニメ作品の最終回には、ツイートユーザ数が上昇しやすいということが経験的に確認されたため、最終回のデータは除外してクラスタリングを行う。次に、各アニメ作品の週間ツイートユーザ数推移データの長さを11にリサンプリングし、週間ツイートユーザ数を

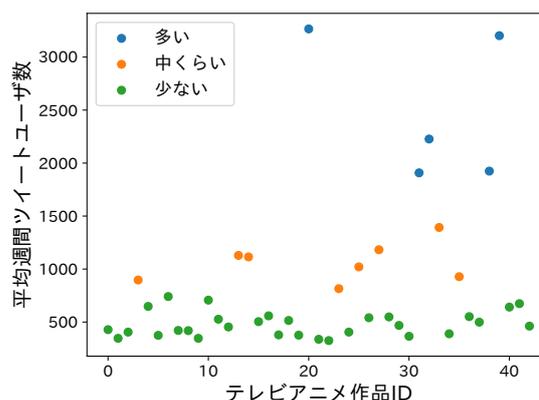


図6 平均週間ツイートユーザ数のクラスタリング結果

0~1の範囲に正規化する。そして、距離尺度をDTWとしてk-means法によりクラスタリングを行った。

クラスタリング結果は図5のようになり、週間ツイートユーザ数の推移傾向として、上昇、下降、山型、横ばいの4クラスタを取得した。作品数はそれぞれ7, 15, 10, 11作品である。

#### 4.3 平均週間ツイートユーザ数のクラスタリング

テレビアニメ作品の平均週間ツイートユーザ数は、上位10%の中でも大きな差があるため、階層的クラスタリング手法の1つであるWard法を用いてクラスタリングを行い、3つのクラスタに分類する。

クラスタリング結果は図6のようになり、平均週間ツイートユーザ数が多い、中くらい、少ない、の3クラスタを取得した。作品数はそれぞれ5, 8, 30作品である。

## 5 感情分析

### 5.1 モデルのファインチューニング

ツイートに含まれる感情を分析するため、感情分析モデルを作成する。モデルは GPT を使用し、事前学習済みモデルとして rinna 株式会社の rinna/japanese-gpt-neox-3.6b<sup>2)</sup> を使用する。ファインチューニング用のデータとして、鈴木らが整備した SNS テキストを用いた感情分析データセットを使用する [5]。このデータセットに含まれている感情極性ラベルは、strong positive, positive, neutral, negative, strong negative の 5 ラベルであるが、これをポジティブ・ネガティブ・ニュートラルの 3 ラベルに単純化して使用する。パラメータはバッチサイズ 64, エポック数 2, 学習率 5e-05 とした。ソースコードは塚越らによって作成された BERT によるテキスト分類のプログラムを参考に作成した [6]。

### 5.2 クラスごとの感情分析の結果

最初に、週間ツイートユーザー数の推移傾向ごとに、オリジナルツイートを対象として感情分析を行なった。その結果は図 7 のようになった。基本的にどの傾向でもポジティブとニュートラルなツイートの割合が高く、ネガティブツイートの割合が極端に低いが、そのなかでも、上昇傾向はポジティブツイートの割合が低く、ネガティブツイートの割合が他の傾向よりも高いことがわかる。逆に、山型傾向はポジティブツイートの割合が高く、ネガティブツイートの割合が他の傾向よりも低いことがわかる。

同様に平均週間ツイートユーザー数によるクラスごとに感情分析を行うと、図 8 のより、基本的にどのクラスでもポジティブとニュートラルなツイートの割合が高く、ネガティブツイートの割合が極端に低いが、平均週間ツイートユーザー数が多いクラスはポジティブツイートの割合が低い。また、中くらいのクラスはポジティブツイートの割合が高く、ネガティブツイートの割合が他のクラスよりも低くなっていることがわかる。

## 6 おわりに

本稿では、繰り返し放送されるテレビアニメ作品を対象に、Twitter 上での情報拡散傾向の違いによるツイートに含まれている感情の違いを分析した。

2) <https://huggingface.co/rinna/japanese-gpt-neox-3.6b>

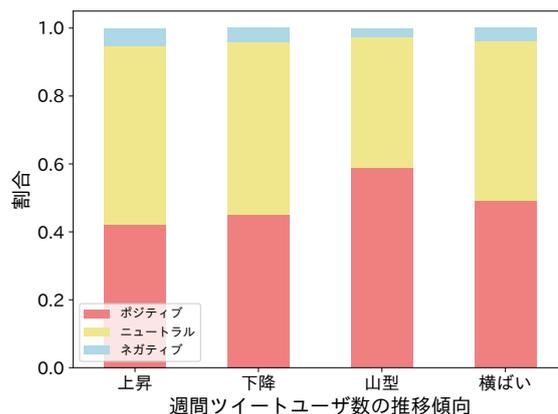


図 7 週間ツイートユーザー数の推移傾向ごとの感情分析結果

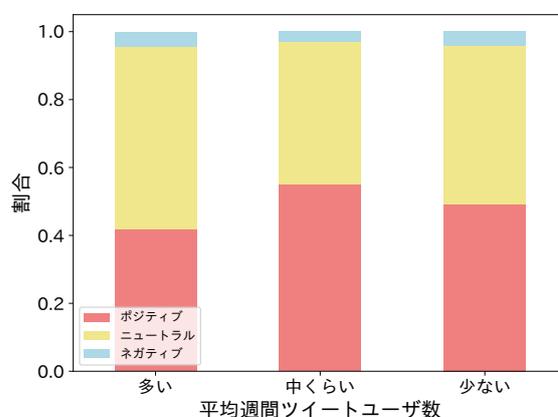


図 8 平均週間ツイートユーザー数によるクラスごとの感情分析結果

情報拡散傾向は、週間ツイートユーザー数の推移データによるクラスタリングと、平均週間ツイートユーザー数によるクラスタリングによって、それぞれ 4 クラスと 3 クラスに分類した。

そして、感情分析モデルを作成し、情報拡散傾向の種類ごとにツイート内容に対して感情分析を行うことで、情報拡散傾向の違いによってツイートに含まれている感情にも違いが現れることを確認した。

今後の課題として、本稿で作成した感情分析モデルは作品のタイトルやハッシュタグに影響されて誤分類していることが懸念されるため、感情分析結果に対し、詳細な評価を行う必要がある。

また、本稿での感情分析では、作品に対する評価のツイートと実況的なツイートを分類できていないため、具体的なユーザー行動を調べることができなかった。したがって、今後はツイート内容をより詳細に分析することで、情報拡散傾向の違いによる具体的なユーザー行動の違いを明らかにしていく。

## 参考文献

- [1] Peihua Fu, Anding Zhu, Qiwen Fang, and Xi Wang. Modeling periodic impulsive effects on online tv series diffusion. **PLOS ONE**, Vol. 11, No. 9, pp. 1–21, 2016.
- [2] 榑剛史, 鳥海不二夫, 大知正直. ソーシャルメディア上の大規模情報拡散に関する俯瞰的可視化手法の提案. 人工知能学会全国大会論文集, 2019.
- [3] Shivangi Chawla and Monica Mehrotra. Impact of emotions in social media content diffusion. **Informatica**, Vol. 45, , 2021.
- [4] 相川実穂, 仙海かおり, 安田智亜紀, 酒井啓江, 小川祐樹, 梅原英一. 深夜アニメ番組に関するつぶやきに注目した twitter 分析. 情報科学技術フォーラム講演論文集, Vol. 14, No. 4, pp. 85–92, 2015.
- [5] Haruya Suzuki, Yuto Miyauchi, Kazuki Akiyama, Tomoyuki Kajiwara, Takashi Ninomiya, Noriko Take-mura, Yuta Nakashima, and Hajime Nagahara. A Japanese dataset for subjective and objective sentiment polarity classification in micro blog domain. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 7022–7028, Marseille, France, 2022. European Language Resources Association.
- [6] 塚越駿, 平子潤. BERT によるテキスト分類チュートリアル. 自然言語処理, Vol. 30, No. 2, pp. 867–873, 2023.

## A 付録

### A.1 感情分析モデルの評価

表 1 は, 先行研究 [5] の感情極性ラベルを単純化したテストデータを用いて, 先行研究の注釈者のアノテーション精度と感情分析モデルの精度を比較した結果である. ファインチューニングにより作成した感情分析モデルは, 注釈者と同等の精度が出ていることがわかり, この感情分析モデルを用いることは妥当だと判断した.

表 1 先行研究の注釈者と感情分析モデルとの精度比較

	正解率	絶対誤差	QWK
注釈者 A	0.802	0.201	0.857
注釈者 B	0.711	0.290	0.745
注釈者 C	0.872	0.129	0.908
感情分析モデル	0.780	0.225	0.828