

放送開始直後に書かれたレビューに注目した アニメ視聴継続の早期判断支援

田村 峻明[†] 莊司 慶行[†] Martin J. Dürst[†]

[†] 青山学院大学 理工学部 情報テクノロジー学科 〒252-5258 神奈川県 相模原市中央区 淵野辺 5-10-1
E-mail: †tamura@sw.it.aoyama.ac.jp, ††{shoji,duerst}@it.aoyama.ac.jp

あらまし 本研究では、Web上のアニメレビューサイトから、アニメの最終的な評判を早期に予測できるレビュー（すなわち、先見の明のあるレビュー）を発見する手法を提案する。現在では年に200本ものアニメが制作されるため、アニメファンはどのアニメを最後まで見続けるか早期に判断する必要がある（俗に「一話切り」と呼ばれる）。アニメレビューサイトにおいて、アニメの放送終了後の平均評価と、あるレビューが放送開始初期につけた評点が類似していた場合、そのレビューは「先見の明のあるレビュー」であると考えられる。SVMを用いてこのようなレビューを分類し、特徴量分析を行うことで、「先見の明のあるレビュー」の特徴を明らかにする。レビューサイト「あにこれ」の実データを用いた実験により、早期にアニメコンテンツの将来価値を推定可能なレビューは放送開始直後に過度に批判的な語を使わない、レビュー歴が長いなどの特徴を明らかにした。

キーワード テレビアニメ, レビューサイト, レビュー分析, 特徴量分析

1 はじめに

近年、テレビアニメは一般的な娯楽として馴染みつつあり、さらに、日本のテレビアニメは海外でも人気を集めている。それに伴い、日本では毎日、多くのテレビアニメが放送されている。ここ数年では、毎年200本¹の新しいテレビアニメが制作されている。ここで、仮にすべてのアニメを一人のアニメファンが1年間にわたって視聴しようとした場合、年間で約1,000時間、すなわち1日のうち約3時間をアニメに費やす計算になる。このような状況下で、全てのアニメを視聴し続けることは時間的に現実的ではない。したがって、年に4回訪れる番組改編期には、アニメファンは新しく始まったアニメを最後まで見るべきかどうかを放送開始直後に（つまり、1話から2話の内容だけ視聴して）判断し、継続的に視聴するアニメを数本に絞る必要がある。この際、キャラクターデザインやアニメ制作会社、声優などの要素から、アニメ放送開始前にあらかじめ視聴するアニメを絞り込むアニメファンも存在する。その一方で、アニメの一話を視聴し、その内容から視聴継続を判断するアニメファンも存在する。このような行為は「一話切り」と呼ばれ、アニメファンの中では一般的に行われている [8], [9]。

しかし、アニメの一話時点で正確にそのアニメが将来的に面白くなるかを推測し、視聴を続けるか判断するのは、困難である。例えば、後半になってから盛り上がる作品や、途中で評価を急落させる作品は数多く存在する。そのため、本来は高く評価されるはずの作品が誤って「一話切り」されてしまうことがある。これは、本来見たかったアニメをリアルタイムで楽しむことができなかったアニメファン本人だけでなく、序盤の内容

だけに基づいて作品全体の質を不当に低く評価される場合など、アニメ制作会社にとっても不利益であり、深刻な問題である。したがって、毎週放送されるテレビアニメにおいて、視聴者が早い時期に正しくアニメを評価できるようにすることは、重要な課題である。

本研究では、Web上のアニメレビューサイトから、アニメの最終的な評判を早期に予測できる「先見の明を持つレビュー」を発見し、そのレビューの特徴を明らかにする手法を提案する。アニメレビューサイトの、あるアニメにおける放送終了後の平均点と、あるレビューが放送開始初期につけた評点が類似していた場合、そのレビューは作品の将来の価値を早期に発見できている、すなわち、先見の明を持つレビューとみなせる。ここで、先見の明を持つレビューと先見の明を持たないレビューを分類するSVM (Support Vector Machine) を作成する。分類に有効だった特徴を分析することで、先見の明を持つレビューと持たないレビューの違いを発見できると考える。

本研究の学術的貢献点として、コンテンツの将来価値を推定するために、コンテンツそのもの（あらすじや制作者などのメタデータ）ではなく、コンテンツの評価者に注目している点が挙げられる。コンテンツの将来価値の推定は、未来予測と同様に、極めて困難な情報学上の課題である。しかし一方で、将来面白くなりそうなアニメを「一話切り」せずに見続ける、アニメの将来価値の予測に長けたアニメファンは確かに存在する。先見の明をもつアニメファンが根強く支持した結果、低い前評判から一転して放送後半や放送終了後に高い評価を得たアニメは「ダークホース」と呼ばれ²、数多く存在する。このようなダークホースアニメを発見できているアニメファンの特徴、す

1: 一般社団法人日本動画協会アニメ産業レポート 2018 サマリー
http://aja.gr.jp/jigyou/chousa/sangyo_toukei

2: ダークホース (アニメ) - アニヲタ Wiki(仮) - アットウィキ
<https://www49.atwiki.jp/aniwotawiki/pages/28788.html>

なわち将来価値の推定がうまいレビューがどこに注目しているか、あるいはどうやってそれを判断しているかを分析することは、未来予測の研究においても重要だと考えられる。

本論文の構成を記す。第2章では、本研究に関連した研究について紹介する。第3章では、提案手法について詳細を説明する。第4章では、実際に有名レビューサイトである「あにこれ」のデータを用いた実験について述べ、第5章で実験結果に対して考察する。第6章では、本研究のまとめと今後の展望を述べる。

2 関連研究

本研究では、先見の明を持つレビューを発見することで、一般ユーザがコンテンツの価値を早期に判断可能にすることを目的としている。先見の明に注目した研究として、先見性に関する先行研究とアーリーアダプターに関する研究について、また将来の人気予測の研究について述べる。加えて、同じくレビューデータを用いた研究として、レビュー分析について述べる。

2.1 先見の明を持つユーザに注目した研究

先見の明に直接言及した別の研究として、掛谷ら [10] は、先見性のある人物と先見性のない人物の特徴を見出す方法として、Amazon のカスタマーレビューを使用した機械学習手法を提案している。掛谷らは、評価が大きく変動した商品のレビューにおいて、多数派の意見に賛成せずにその商品の後の評価につながる評点をつけたユーザには先見性があり、多数派の意見に賛成した評点をつけたユーザには先見性がないと定義している。そして、Amazon の書籍商品のカスタマーレビューを利用し、先見性のあるユーザと先見性のない人物を発見する。その後、最大エントロピー法を用いた機械学習を用い、先見性のあるユーザは洋書を含む本のレビューの割合が多いこと、本のジャンルでは文学・評論作品のレビューが多いことを明らかにしている。また、先見性のないユーザはメディアに流されるレビューをする傾向があることを明らかにしている。本研究では、これと異なり、複数の話にわたる「アニメ」について先見性のあるユーザに関しての研究をしている。

先見の明のあるレビューは、物の価値に早期に気づく能力を持ったユーザであると見做すことができる。このようなユーザはマーケティングなどの分野で「アーリーアダプター」と呼ばれており、情報学分野でもその性質に注目した研究が一般的に行われている。

今森ら [4], [11] は、Twitter において、他者に先駆けて優良な情報源を発見する能力に優れたユーザをアーリーアダプターと定義し、それに注目することで、新規情報源の将来の人気をランク付けする手法を提案している。アーリーアダプターの発見は、Twitter におけるフォロー、フォロワー関係を参照し、フォローを模倣されている数に注目して推定している。

Tyler ら [7] は、イノベーターと「のろま (Laggard)」を分類するためにオンライン製品レビュー Web サイト「Epinions.com」のユーザの注文の統計を作成し、ユーザの採用パ

ターンを分析している。その結果、Tyler らはそれぞれの製品のレビューから、カテゴリ合計のランキングがそれぞれの製品をレビューするユーザの作成したランキングと一致することを実証している。

先見の明を持つレビューは、アーリーアダプターや先見性のある人物、イノベーターなどの呼ばれ方をしているが、コンテンツの性質を把握する用途で注目を集めていることがわかる。

2.2 人気予測に関する研究

あるアイテムが将来的にどのような評判を得るかを推測する問題は、情報学だけでなく経済やマーケティングなど、幅広い分野で盛んにおこなわれている。

Pinto ら [6] は、あるコンテンツに対して、特定の参照日までの、日々の人気度のデータを利用し、コンテンツの将来の人気を予測するための手法を提案している。Web 動画サイト「YouTube」のデータを用いて実験を行っている。特定の参照日までのコンテンツの人気度の推移に、それぞれ日ごとに異なる重みを割り当てる。これにより、様々な人気の発展パターンを持つ動画を、より正確に区別することを可能にしている。Pinto らの今後の展望として、他の変数、例えばユーザ関連の変数の導入が予測精度を改善するかどうかを調査することを挙げている。

Ouyang ら [5] は、様々な特徴と有効的な分類手法を用いて、オンラインビデオの将来の人気レベルを予測している。中国の大手動画サイト「Youku」のデータを用いている。人気レベルの遷移に基づく、将来の再生数を予測する回帰モデルを構築し、さらに分析により、過去の人気度に関する特徴が予測の精度に大きく関わることを明らかにしている。これらの研究のように、コンテンツの内容やメタデータから、コンテンツの将来の価値を予測する研究は一般的に行われている。一方で、本研究ではコンテンツではなく、ユーザ情報に着目しコンテンツの早期予測を行うことを大きな目的としている。

2.3 オンラインレビューの分析

オンラインレビューに着目し、レビューが与える影響を分析する研究は、情報学、経済学、マーケティングなどの幅広い分野で行われている。Hao ら [2] は、ネットのレビューのポジティブな意見やネガティブな意見が、消費者の意思決定において重要かどうか分析を行っている。帰属理論と見込み理論の2つの手法を利用し検証しており、探索的属性を持つ商品に関して、ポジティブレビューとネガティブレビュー間で口コミの効果に違いはないが、経験的属性を持つ商品に関しては、ネガティブレビューは口コミ効果が大きいと述べている。本研究で対象とするアニメ作品は、実際に見てみなければわからないコンテンツであるため、経験的属性を持つ商品に近い。そのためネガティブなレビューがより読者に影響を与えることが考えられる。ここから、特徴に批判的な語に関する項目を加えた。

レビューによる将来価値の推定を行っている研究例として、野中ら [12] は、レビューにおける消費者の感情を分類し、予測タスクへの寄与度の高い特徴を用いることで、販売予測のタス

クの精度を向上する手法を提案している。実験の結果、精度が向上していることから、レビューは商品の売りに影響があることを実証している。

ほかに Connors ら [1] は、レビューのどの要素が意思決定に有用であるかについて研究している。レビュー内容とレビューの特性の両方が、レビューの有用性を決定づけていることを明らかにしている。本研究では、レビュー中の要素でなく、人に注目して、先見の明を持つレビューの早期レビューは一話切りの際の意思決定に有用であると考えており、既存のレビューからの意思決定支援の研究と異なる。

3 SVMを用いた「先見の明のあるレビュー」分析

本研究では、Web上のアニメレビューサイトから、アニメの最終的な評判を早期に予測できるレビュー、すなわち先見の明を持つレビューを発見し、その先見の明を持つレビューがどのような特徴を持っているのかを明らかにする。

本手法では、はじめにアニメレビューサイトに対してクロールリングを行い、分析対象とするデータを収集する。収集したデータから、早期にアニメを正しく評価できていたレビューを抽出する。こうしたレビューの性質について様々な仮説をたて、レビューを特徴ベクトルで表す。次に、その特徴ベクトルから、そのレビューが先見の明を持つか持たないか SVM (Support Vector Machine) で分類する。最後に、交差検証を通して分類に大きく寄与した特徴量を発見する。分類に大きく寄与した特徴に注目して、それぞれのレビューを分析することで、先見の明を持つレビューの特徴を明らかにする。

3.1 アニメレビューサイトのクロールリング

アニメレビューサイトから、分析対象のデータを収集する。アニメレビューサイトは、通常のレビューサイトと同様に、「アイテムのメタデータ」、「レビューのプロファイル」、「あるレビューからあるアイテムに対するレビュー」の3つの要素から構成される。

「アニメ作品のメタデータ」は、「アニメの放送期間」や「アニメ ID」、「アニメの平均評点」などを含む。アニメレビューサイトにおいて、アニメは独自の ID で管理されている。各レビューがどのアニメについてレビューをしたかを結びつけるのに用いるため、取得する。また、「アニメの放送期間」も重要である。本実験では、頻繁に「一話切り」の対象となる1クール(12話前後)または2クール(24話前後)のアニメに焦点を絞って分析を行う。アニメの放送期間は、あるレビューがアニメの放送開始から、どの時点でつけられたかを確認するために用いられる。各アニメについて、アニメにつけられた全てのレビュー評点の平均である「アニメの平均評点」も取得する。これは、レビューが正確にアニメを評価できているか確認する際に用いる。

「レビューのプロフィール」は、それぞれのレビューに関する任意入力のプロフィール情報と、フォロー関係などのソーシャル要素、レビュー本数などの統計情報を含む。レビューとレビ

ュアを結びつけるために、「レビュー ID」を取得する。レビューのレビューサイトへの熱意を確認するために、「プロフィールコメント」を取得する。また、「アニメ棚のアニメ総数」も取得する。アニメ棚とは、レビューがアニメ情報を確認するために、任意のアニメをリスト化して管理できるシステムである。これは、そのリストに入れられたアニメ本数のことである。他に、「レビュー総数」、「サンキュー数」、「フォロー数」、「フォロワー数」、「メッセージ数」もそれぞれ取得する。

「あるレビューからあるアニメに対するレビュー」は、個々のレビューの内容そのものを含むデータである。それぞれのレビューに振り分けられた、レビューを特定するための「レビュー ID」を取得する。どのレビューがそのアニメをつけたかを確認するために必要である。レビューの特性を取得するために、「レビュー本文」と「レビュータイトル」を取得する。他に「アニメ ID」、「レビュー日時」、「レビュー評点」もそれぞれ取得する。

3.2 先見の明を持つレビューの抽出

本実験における先見の明を持つレビューは、アニメの最終的な評判をアニメ放送開初期に予測できるレビューを指している。

先見の明を持つレビューを定義し、抽出する。本研究において分析に用いる先見の明を持つレビューの条件として、

- (1) テキスト付レビューを投稿している、
- (2) 早期レビューを5件以上投稿している、
- (3) 早期レビューの評点そのアニメの最終評価と近いの3件を設定した。この際、「放送開始1か月以内に投稿されたレビュー」を「早期レビュー」と呼ぶ。

テキスト付レビューの有無について、これはクロール対象としたアニメサイトが、評点だけ入力してテキストが空の状態のレビューを受け付けているためである。本研究ではあるレビューを表す特徴としてレビューの文体などを用いるため、一度もテキスト付レビューを投稿していないレビューを除外した。早期レビューの投稿数について、ある程度の本数がないと先見の明があるかどうかの判断が出来ず(すなわち、たまたま1本から2本のアニメについて将来価値を正しく推定した可能性がある)、またレビューの早期レビューにおける特徴の分析も困難になるためである。早期レビューとアニメの最終評点との比較について、差の大きさを順位付けて、上位のレビューを先見の明があると見做す。

早期レビューの定義をアニメ放送開始から1ヶ月以内と決定した理由として、実用面での理由とデータ量による理由がある。1ヶ月という期間は、毎週決まった時間に放送されるアニメにおいて、最大で4話(すなわち、全放送の3分の1)が放送されている時期である。コンテンツの3分の1だけを視聴してアニメの最終評価を推定できるレビューは、本研究における分析を行う上では、十分に先見の明を持つと判断したためである。また、データ量の問題として、そもそも放送開始最初期にレビューを書くユーザが多くなく、地域によって放送開始時期が異なるため、厳密に1週間だけのデータを用いた分析が現実的でなかったためである。

以上の定義に基づき、先見の明を持つレビューを発見する手順を述べる。全てのレビューに対し、レビュー日時と、レビュー対象のアニメの放送開始時期を比較し、レビュー日時とアニメの放送開始時期1ヶ月以内のレビュー、すなわち早期レビューをリストアップする。リストアップした早期レビューから、5件以上レビューを書いたレビューを早期レビュワー者として抽出する。早期レビュワー者から、本実験の対象外であるレビュー文を書いたことのないレビューを除外する。上記の早期レビュワー者が書いた早期レビューの評点と、アニメレビューサイトに掲載されているアニメの放送終了後の評点の平均点を比較し、その誤差を求める。

誤差の指標として、誤差の二乗の平均を用いる。あるレビューア r の n 個めの早期レビューの評点を x_n 、その早期レビューの対象のアニメの最終平均点を y_n とする。この際、誤差 $d(r)$ は、

$$d(r) = \frac{1}{n} \sum_{i=1}^n (x_n - y_n)^2 \quad (1)$$

と表される。

全てのレビューアについて計算した $d(r)$ を比較し、 $d(r)$ が小さいレビューアを先見の明を持つレビューア、 $d(r)$ が大きいレビューアを先見の明を持たないレビューアとして分析に用いる。

3.3 特徴量作成

上記のデータを用い、先見の明を持つレビューアの特徴の仮説をたて、特徴量を作成する。特徴量を作成するうえで注目する要素として、レビューアの熱意、アニメ知識、ソーシャル関係、レビューの着眼点、レビューしたアニメの放送年、レビュータイミングを挙げる。

レビューアの熱意は、レビューアが、レビューを書いたアニメに対して、もしくはレビューサイトに対して、どれだけ執着しているかを表す。レビューしたアニメに対して執着があるレビューアほど、そのアニメを深く理解し、より考察していると考えられ、アニメを正しく評価ができると考えられる。熱量に分類される特徴として、文字数が多い、感情的な表現の使用などが考えられる。これを表すための具体的な特徴量として、レビューの文字数、レビュー本文内の感情語の使用頻度などを用いる。

アニメ知識は、そのレビューアがアニメに関する知識をどれだけ持っているかを表す。アニメ知識を持っているほど、様々なアニメに精通しており、アニメを正しく評価ができると考えられる。アニメ知識に関する特徴として、チェックしているアニメ本数が多い、アニメ作成に携わる人物を認知しているなどが考えられる。これを表すための具体的な特徴量として、アニメ棚（ウォッチリスト）に含まれるアニメ本数、これまでに書いたレビューの総数、レビュー本文における声優名の使用頻度などを用いる。

ソーシャル関係は、他のレビューアとどれだけ関係性を持っているかを表す。ソーシャル関係に分類される具体的な特徴量として、獲得サンキユー数、フォロー数などが挙げられ、それらの特徴量として用いる。サンキユー数が多いことは、他のユーザから信頼を得ている、また的を射たレビューをつけているこ

とを表すと考えられる。フォロー数が多いことは、アニメの知識を得ることに貪欲であると言えるため、正しい評価を行える可能性が比較的高いと考えられる。

レビューの着眼点は、レビューがアニメのどこに注目しているかを表す。先見の明を持つレビューアは、先見の明を持たないレビューアと比べて注目している要素が違うのではないかと考えられる。これを表すための具体的な特徴量として、ストーリー展開、キャラクターに触れているかなどを用いる。具体的には、ある1人のレビューアが書いたレビューについて、ストーリー展開は「テンポ」、「構成」などの単語をレビュー文に含む割合である。同様にキャラクターは「可愛い」、「カッコいい」などの単語をレビュー文に含む割合を用いる。

レビューしたアニメの放送年は、レビューアがどの年のアニメレビューを多く投稿しているかを表す。昔のアニメを見ている割合が他のレビューアと比べて大きいレビューアほど、アニメレビューの経験が長く、アニメを正しく評価ができると考えられる。これを表すための具体的な値として、各年のアニメレビューの割合を用いる。

レビュー時期は、レビューアが、放送開始からどのくらい期間をおいてレビューを書いたかを表す。放送開始直後にレビューを書いている割合が大きいレビューアほど、アニメのその後を予測することに慣れていていると考えられる。これを表すための具体的な特徴量として、放送1ヶ月以内のレビュー割合などを用いる。

これらの仮説から、任意のレビューアを表す計76次元の特徴ベクトルを作成する。主な特徴を表1に示す。

3.4 交差検証による特徴量分析

先見の明を持つレビューアとそうでないレビューアの特徴の違いを明らかにするためにSVMを利用する。SVMは、線形入力素子を利用し、2クラスに識別する分類機である。本研究は、先見の明の有無によりレビューアを2種類に分類する。多次元の特徴に強く、2クラス分類を行うのに適したSVMを用いる。

あるレビューアを表す特徴ベクトルを入力とし、先見の明を持つレビューアか分類する。交差検証を行うことで、作成した分類機の精度を評価する。ここで、特徴量分析のため、はじめに全ての特徴量を入力とし、分類精度を計算する。そして、特定の特徴を1次元除外したものを入力として、同様に分類精度を計算する。前者と後者を比較した際に、精度が大きく下がった場合は、除外した特徴量が分類に大きく寄与したと考えられるため、その特徴は先見の明を持つレビューアの特徴であると判断できる。

4 実験

提案した手法に基づいて、「先見の明のあるレビューア」について、アニメレビューサイト「あにこれ」³のデータを用いた分析を行った。本章では、実際に収集したデータから、先見の明

3: おすすめアニメ動画を感想評価/人気でランキング【あにこれβ】

<https://www.anikore.jp/>

表 1 先見の明の有無の推定に用いる、レビューを表す特徴の例（合計 76 次元）

注目する点	特徴	実際の値
レビュー文の着眼点	展開に言及することが多い キャラクターに言及することが多い 原作前作に言及することが多い	「構成」,「王道」などの語を含む割合 「可愛い」,「かっこいい」などの語を含む割合 「原作」,「前作」などの語を含む割合
レビューアニメ	最新アニメを見ている 昔のアニメを見ている	2018 年アニメのレビューの割合 2008 年アニメのレビューの割合
レビュー時期	放送終了直後のレビューが多い 放送開始直後のレビューが多い	放送終了一ヶ月以内のレビューの割合 放送開始一週間以内のレビューの割合
レビューの熱意	長いレビューを書く 感情的なレビューを書く	レビューの文字数 感情語の使用割合
アニメ知識	たくさんアニメを見る 見たいアニメが多い 声優に言及することが多い	投稿したレビューの数 アニメ棚のアニメの本数 「山寺宏一」などの声優名を含む割合
ソーシャル関係	サンキュー数が多い フォロー数が多い フォロワー数が多い メッセージ数が多い	サンキュー数 フォロー数 フォロワー数 メッセージ数

を持つレビューを抽出した。その後、先見の明を持つレビューを特徴ベクトルで表し、SVM を用いて実際に先見の明を持つレビューと先見の明を持たないレビューを分類をした。

4.1 使用したデータセット

本研究で使用したデータセットの詳細について述べる。今回実験に使用したデータは、アニメレビューサイト「あにこれ」から収集した。クロウリング手法として、Ruby と Nokogiri を用いてスクレイピングを行った。2018 年 9 月 6 日から 2019 年 1 月 11 日にかけてデータを収集した。

データの収集元としてアニメレビューサイト「あにこれ」を選んだ理由として、アニメに関するレビューサイトの中で、最もデータの規模が大きかったことを挙げる。表 2 に主要なアニメレビューサイトの一覧とそのデータ規模を示す。

2008 年春から 2018 年春までに放送されたアニメ作品について、それぞれのアニメ作品のメタデータを 2,307 件、これらのアニメに対するレビューを 987,964 件、これらのレビューを投稿したレビューアのプロファイルを 47,958 件、それぞれ収集した。実際の実験では、データが不十分なものと 2 クールより長いアニメへのレビューを除外し、790,219 件のレビューを使用した。作成したデータセットに含まれる情報の一覧を表 3 に示す。

始めに、本研究で対象としないアニメをデータセットから取り除いた。本実験で対象とするアニメは、商業用のアニメーション作品のうち、テレビで放送された連載形式のものを指す。連載アニメの放送期間は季節ごとに区切られており、1 つの季節（12 話から 13 話）を 1 クールと呼ぶ。本研究では新規アニメにおける視聴継続の早期判断支援をテーマに設定しているため、1 クールまたは 2 クールのアニメを対象とする。そのため、長期にわたって放送されるアニメや 1 話のみ放送されたアニメを除外した。

全てのレビューに対し、レビュー日時とアニメ放送時期を比

表 2 各レビューサイトに登録されたアニメ本数と、最もレビューを多く獲得したアニメについてのレビューの件数

アニメレビューサイト	アニメ件数	レビューの最大数
あにこれ	11,975	15,417
アキバ総研	6,018	235
アニメ評価データベースさち	5,468	405
作品データベース	7,189	2,672

較し、レビューが放送開始からどのタイミングで書かれたかを分析した。書かれた時期とレビュー件数を表 4 に示す。これより、アニメ放送開始直後に書かれたレビューは非常に少なく、8 割以上のレビューはアニメ放送終了後に書かれていることが分かった。

注意すべき事項として、今回データセットとして用いた「あにこれ」では、あるレビューアがあるアニメ作品に対して投稿したレビューを更新可能であった。そのため、早期にレビューを投稿していたとしても、あとからその評価を上書きした場合、早期レビューとみなされず、最終評価に反映される。

ここで、収集したデータセットにおけるレビュー評点の分布を、図 1 に示す。多くのレビューサイトにおいて、レビューが絶賛と酷評に二分化して、2 点から 4 点が少ないという現象が指摘されている [3]。本実験のデータセットについても同様のことが予想されたが、概ね正規分布に従う傾向が見られた。そのため、本実験においては、先見の明を持つかの正解として、放送終了後の平均点を用いる。ここで、入力時の初期値である 3 に評価が集中している現象がみられる。この原因は、本文にアニメへの感想だけを書きたいレビューが多く、その際に初期値である 3 のまま投稿していることが考えられる。

4.2 先見の明を持つレビューの抽出

上記のデータセットから、実際に先見の明を持つレビューを抽出した。放送開始 1 ヶ月以内のアニメに対し、レビューを 5

表 3 実験に使用したそれぞれのデータセットのフィールド

レビュー	プロフィール	アニメ情報
レビュー ID	レビュー ID	アニメ ID
レビュー名	レビュー名	アニメタイトル
レビュー内容	自己紹介文	アニメ放送期間
アニメ ID	アニメ棚の本数	アニメ平均点
レビュー日時	レビュー総数	
レビュー評点	獲得サンキュー数	
	フォロー数	
	フォロワー数	
	メッセージ数	

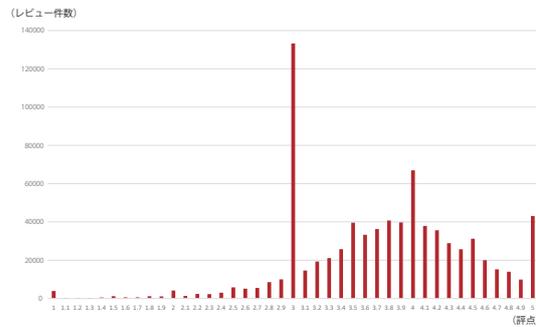


図 1 評点ごとのレビューの件数

表 4 放送開始からレビュー投稿までの期間ごとのレビュー件数の推移

書かれた時期	レビュー件数
放送開始 1 週間以内	3,844
放送開始 2 週間以内	1,235
放送開始 3 週間以内	17,334
放送開始 1ヶ月以内	24,892
放送開始 2ヶ月以内	43,234
放送開始 3ヶ月以内	53,185
放送中	112,235
放送終了後	677,984
全レビュー数	790,219

件以上書いた早期レビューは 1,233 人存在した。そのうちレビューコメントを書いていないレビューを除外し、残った早期レビュー者は 853 人となった。そのレビューに対し、式 1 を用い誤差 $d(r)$ を計算した。

全ての早期レビュワーについて $d(r)$ の小ささにおいて上位 100 人を先見の明を持つレビューア、下位 100 人を先見の明を持たないレビューアとして用いる。それぞれの $d(r)$ の閾値を表 5 に示す。

先見の明を持たないレビューアの $d(r)$ の最小値は、先見の明を持つレビューアの $d(r)$ の最大値の約 5 倍ある。ここから、アニメの将来の価値を推定する能力は人によって大きく異なることが分かる。

4.3 先見の明を持つレビューアの特徴量作成

提案手法に基づき、取得したデータセットを用いて、先見の明を持つレビューアの特徴量を計算し、それぞれのレビューアを特徴ベクトルで表す。表 1 にあるように、それぞれの特徴の実際の値を計算する。「レビュー文の着眼点」の要素は、レビューがアニメのどこに注目しているかを表している。これらの特徴は、人手で作成した辞書内の語を含むレビューの割合によって表される。レビューがアニメの展開に言及することが多かった場合には、具体的な特徴量として、レビュー中に「展開」、「構成」などの単語を含んでいた割合を用いる。他シリーズに言及することの多さを表すためには、「原作」、「前作」、「ゲーム」などの単語をレビュー文に含んでいた割合、アニメのジャンルについて言及することが多かった場合には、「学園」「ギャグ」などの単語の割合を用いた。また、声優とアニメスタッフの一覧を作

表 5 早期の評点と放送終了後の平均評点の差 $d(r)$ の閾値

	最小値	最大値
先見の明を持つレビューア	0.012	0.18
先見の明を持たないレビューア	0.89	6.25

成し、これらへの言及割合も用いた。

4.4 SVM を用いたレビューアの分類

SVM で先見の明を持つレビューアと先見の明を持たないレビューアを実際に分類した。SVM の実装として Python の scikit-learn を用いた。与えられた特徴量から、先見の明を持つレビューアを 1、先見の明を持たないレビューアを 0 と出力するように学習させる。入力する特徴量は、scikit-learn に搭載されている、StandardScaler を用いて標準化した。交差検証には CrossValScore を用い、5 交差検証で精度 (accuracy) を算出した。すなわち、200 名の先見の明を持つレビューアと持たないレビューアについてデータを 5 分割し、40 名ずつのサンプルを作成した。1 つをテストデータ、残り 4 つを訓練データとして利用し、5 回の分類の精度の平均を最終精度として用いる。

分析を始める前に、最も精度の高い状態の分類器を作成する。用意した全ての特徴量をそのまま用いた場合、分類精度が 0.63 と低い値をとった。これは、特徴量のうちいくつかが先見の明を持つか持たないかに依存するのではなく、個人個人に差がある特徴量であり、その特徴量が分類に悪影響を与えているためだと考えられる。そこで、分類に不必要だと考えられる特徴量を入力から除外した。具体的な例として、キャラクターへの言及割合という特徴量は、キャラクターの捉え方というのは、アニメ全体の評価につながる要素よりも、個人個人の趣味趣向が表れる可能性が考えられる。これらの特徴は、分析上障害となるため除外した。

上記のように特徴量を選別し、計 58 次元の特徴量を SVM に入力として利用することで 0.7 の分類精度が得られた。

4.5 特徴量分析

どの特徴量が分類に有効だったかを Feature Ablation と呼ばれる手法で分析した。最も精度が高くなったモデルに対して、ある任意の特徴量を 1 次元分だけ入力から除去した際に、分類精度がどの程度下がるかで、除外した特徴量がどれだけ正しい

表 6 各特徴を除外した際の分類精度

除外した特徴	平均分類精度
何も除外しない場合	0.700
早期レビューでの批判的な単語の使用割合	0.655
2012 年のアニメへのレビューの割合	0.665
2014 年のアニメへのレビューの割合	0.675
早期レビュー以外での同シリーズの作品に関する単語の使用割合	0.685
アニメ開始一週間以内のレビュー数	0.685
アニメ放送終了直後のレビューの割合	0.685
早期レビューの平均文字数	0.685
2013 年のアニメへのレビューの割合	0.685
早期レビューでのストーリーの新規性に言及する単語の使用割合	0.685
早期レビュー以外での批判的な単語の使用割合	0.685

分類に寄与するか計算する。

特定の特徴を除外して分類した際、下がった精度が大きかった上位 10 個の特徴と分類の精度を表 6 に示す。レビューが先見の明を持っているか持っていないかは、これらの特徴に表れていることが分かった。

精度に大きく影響した特徴から、先見の明を持つレビューは放送開始直後に書かれたレビューにおいて、過度に批判的な語を使わない傾向があることが分かった。具体的には、「クソ」、「無駄」、「最悪」などの語を含むレビューが少なかった。一方で、これらのレビューは、放送開始一ヶ月後以降のレビューでは批判的な語を使う傾向もあることが分かった。また、先見の明を持つレビューはアニメ開始一週間以内のレビュー数が少ないこと、日頃のレビューにおいて「原作」、「前作」などの単語を用いて、同シリーズの作品に言及する傾向があること、早期レビューでストーリーの新規性に言及している傾向があること、早期レビューの平均文字数が多いことが、それぞれ分かった。また、これまでに各年のアニメをどれだけレビューしてきたかについて、2012 年から 2014 年にかけてのレビュー経験が分類精度に寄与することが分かった。

5 考 察

本章では分特徴量分析から、先見の明を持つレビュー、また先見の明を持たないレビューがどのような特徴があるかを考察する。

分類精度に大きく影響した特徴から、先見の明を持つレビューは放送開始直後には、レビューにおいて強く批判的な語を用いない一方で、放送後半以降では用いていることが分かった。これは、アニメ放送開始直後に罵詈雑言を含むレビューを投稿して、そのあとの展開についてはレビューを投稿しないレビューは、そもそも信頼できないレビュー（荒らし行為を行うユーザや、「アンチ」と呼ばれるユーザ）である可能性を示唆している。

さらに、先見の明を持つレビューは、日頃のレビューにおいて、「原作」、「前作」などの、そのアニメと関連するシリーズ作品の別作品に関する単語を用いている。このような言及を行うため、アニメ化された作品の原作となった漫画や小説、また過去のシリーズ作品に関する知識も持っている必要がある。このことから、先見の明を持つレビューは、レビュー対象となった

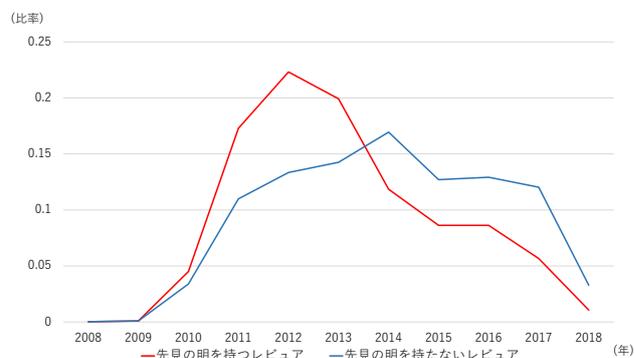


図 2 レビューしたアニメの放送年の割合

アニメそのものだけでなく、関連する幅広い分野に興味を持っている可能性が示唆される。

ここで、分類精度に「レビューしたアニメの放送年」の要素が分類に大きく寄与しているため、先見の明を持つレビューと持たないレビューにおける、2008 年から 2018 年のアニメへのレビュー割合の平均をそれぞれ図 2 に示す。2008 年、2009 年のアニメへのレビューはそもそも絶対数が少ないため比較できないが、先見の明を持つレビューは 2010 年から 2013 年までのアニメへのレビュー割合が高いことが分かる。一方で、2014 年以降へのアニメへのレビュー割合は、先見の明を持たないレビューの方が高いことが分かる。これは、長期にわたってアニメをレビューしてきた、経験の多いレビューが先見の明を持つ可能性が高いことを示唆している。

また、早期レビューの平均文字数が分類精度に大きく影響したため、先見の明を持つレビューは、早期レビューにおいて長文のレビューを投稿する傾向があることが分かった。実際の分類に使用した特徴量の値から、早期レビューの平均文字数を分析する。先見の明を持つレビューとそうでないレビュー 200 人を、早期レビューの平均文字数順に並べると、上位 10 パーセントは 80 パーセント、上位 5 パーセントは 90 パーセントが先見の明を持つレビューであった。同様に、早期以外のレビューにおいても、平均文字数について比較すると、上位 10 パーセントは 75 パーセント、上位 5 パーセントは 80 パーセントが先見の明を持つレビューであった。したがって、時期にかかわらず、レビューの文字数が特に多いレビューが、先見の明を持つ可能性が高いことが分かった。

一方で、分類精度にほとんど寄与しなかった特徴に目を向けると、「フォロー数」、「フォロワー数」、「サンキュー数」などの、ソーシャル関係に属する特徴は除去しても分類精度が下がらなかった。このことから、アニメレビューサイトにおける他者との関係性の有無は、先見の明を持つかに関係しないと考えられる。

本研究における分析全体に対する考察として、本実験において精度が低くなった理由として、利用できる早期レビューのデータ数が少なかったことが考えられる。データ数が少なかった第一の理由として、放送開始直後のアニメに対してレビューするレビューが、全レビューに対してごく少数である可能性が考えられる。多くのレビューは、アニメを全て見終わった時に、

初めてアニメのレビューする機会が多いと考えられる。次に、
 あにこれの仕様上の問題が挙げられる。あにこれでは、あるア
 ニメに対して既にレビューを書いていた場合、次にまた同じア
 ニメにレビューを書こうとすると以前のレビューが消える仕様
 になっている。つまりあるレビューが早期にレビューをつけて
 いたとしても、上書きされている可能性があるため、早期レ
 ビューが少なくなった。

また、特徴量作成の際、人手で辞書を作成したが、それぞれ
 の辞書の語彙数が少なく、レビュー中の単語を網羅しきれなかつ
 た可能性が考えられる。各分野の辞書、すなわちストーリーの
 展開に関する語や、罵詈雑言に分類された批判的な語をより詳
 細な辞書で分類して特徴量として用いることで、分類ならびに
 分析の精度を向上できる可能性がある。

6 まとめと今後の課題

本研究では、Web 上のアニメレビューサイトのデータを用
 い、アニメの最終的な評判を早期に予測できるレビュー、すな
 わち先見の明を持つレビューを発見する手法を提案した。アニ
 メレビューサイト「あにこれ」のデータを用いて実験した結果、
 先見の明を持つレビューは、アニメレビュー歴が長いことが分
 かった。一方、先見の明を持たないレビューは、放送開始直後
 のレビューにおいて批判的な単語を使用する傾向があることが
 分かった。

精度が上がらなかった理由として、データが少ないことを挙げ
 たが、解決策としてアニメレビューサイトだけでなく、Twitter
 等の SNS を利用し分析することが考えられる。SNS には、レ
 ビューサイトより大量のユーザが存在しており、頻繁に新規ア
 ニメに関する話題が取り上げられている。加えて、ユーザのフォ
 ロー関係や過去の発言等も用いることができるため、レビュー
 サイト単体よりもより情報量の多いデータを利用できると考え
 られる。

本研究の知見を活かしたアプリケーションの開発も今後の展
 望として重要である。本研究では、アニメの「一話切り」を支
 援することで、アニメ業界に貢献することが目標であった。そ
 のために、正しい「一話切り」を支援する意思決定支援のため
 のアプリケーションを開発することが今後の展望として挙げられ
 る。本研究で得られた先見の明を持つレビューの特徴を利用し、
 そういったレビューがどこに注目しているかを示すことで、ア
 ニメ視聴者の一話切りにかけるコストを低減すると同時に、ア
 ニメ視聴初心者による誤った一話切りも防げるようになると考
 えられる。

謝 辞

本研究は JSPS 科研費 18K18161 (代表: 莊司慶行), 18H03243
 (代表: 田中克己) の助成を受けたものです。ここに記して謝意
 を表します。

- [1] Laura Connors, Susan M Mudambi, and David Schuff. Is it the review or the reviewer? a multi-method approach to determine the antecedents of online review helpfulness. In *44th Hawaii International Conference on System Sciences (HICSS2011)*, pp. 1–10. IEEE, 2011.
- [2] YuanYuan Hao, Qiang Ye, YiJun Li, and Zhuo Cheng. How does the valence of online consumer reviews matter in consumer decision making? differences between search goods and experience goods. In *Proceedings of 43rd hawaii international conference on System sciences (HICSS2010)*, pp. 1–10. IEEE, 2010.
- [3] Nan Hu, Paul A Pavlou, and Jennifer Zhang. Can Online Reviews Reveal a Product’s True Quality?: Empirical Findings and Analytical Modeling of Online Word-of-Mouth Communication. In *Proceedings of the 7th ACM conference on Electronic commerce*, pp. 324–330. ACM, 2006.
- [4] Daichi Imamori and Keishi Tajima. Predicting popularity of twitter accounts through the discovery of link-propagating early adopters. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 639–648. ACM, 2016.
- [5] Shuxin Ouyang, Chenyu Li, and Xueming Li. A peek into the future: Predicting the popularity of online videos. *IEEE Access*, Vol. 4, pp. 3026–3033, 2016.
- [6] Henrique Pinto, Jussara M Almeida, and Marcos A Gonçalves. Using early view patterns to predict the popularity of youtube videos. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 365–374. ACM, 2013.
- [7] Sarah K Tyler, Shenghuo Zhu, Yun Chi, and Yi Zhang. Ordering innovators and laggards for product categorization and recommendation. In *Proceedings of the third ACM conference on Recommender systems*, pp. 29–36. ACM, 2009.
- [8] アニメ DVD・BD の売り上げを見守るスレ@ wiki. 用語集/1話切り. <https://www24.atwiki.jp/anime-urisure/pages/487.html>.
- [9] ニコニコ大百科. 第〇話切りとは (ダイマルワギリとは) [単語記事]. <https://dic.nicovideo.jp/a/第〇話切り>.
- [10] 英紀掛谷, 裕也佐藤. 書籍のレビューに基づく先見性のある人物の特徴分析. 言語処理学会 第 22 回年次大会 発表論文集, Mar 2016.
- [11] 今森大地, 田島敬史. Twitter におけるアーリーアダプター推定手法の評価. 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 (DEIM2016), G6-2, 2016.
- [12] 野中尚輝, 中山浩太郎, 松尾豊. オンラインレビューから抽出した消費者の感情に寄与する素性を用いた自動車販売予測. 情報処理学会論文誌データベース (TOD), Vol. 10, No. 3, pp. 16–25, 2017.