

ニコニコ動画のログデータを用いた ソーシャルノベルティのある動画の発見に関する研究

平澤真大[†] 小川祐樹^{††} 諏訪博彦[†] 太田敏澄[†]

インターネットの普及によって、ニコニコ動画のような動画共有サイトの需要が高まり、結果多くの動画コンテンツが蓄積されている。これら蓄積された動画コンテンツの中には多くの人には知られていないが、視聴した際に多くの人の興味・関心が湧くコンテンツが多く埋もれていると考える。我々はソーシャルノベルティのある動画を「社会的には知られていないが、より多くの人が興味・関心を持つコンテンツ」と定義し、ソーシャルノベルティのある動画を発見するため「もっと評価されるべき」タグに注目した。本稿ではソーシャルノベルティのある動画発見のため、「もっと評価されるべき」タグの分析と、それを用いた機械学習の精度分析の結果を報告する。

Finding Social-Novelty videos sunk in NicoNico Douga

Masahiro Hirasawa,[†] Yuki Ogawa,^{††}
Hirohiko Suwa[†] and Toshizumi Ohta[†]

Recently, the Web technology enables people to watch VoD services such as NicoNico Douga. However, there are so many video contents in the site that Not known but interesting, “Social-Novelty” videos are sunk in them. To support users to find interesting video sunk in NicoNico Douga, a recommendation system of “Social-Novelty” videos is proposed. Tags of contents rated as “would be higher rated” are analyzed, and found typical tags attached to such contents. This paper reports the result of analysis of the tag “would be higher rated”, and precession analysis of a classification of the machine learning using it for the recommendation system with “would be higher rated”.

1. はじめに

近年、インターネットの普及によって大容量の動画を誰でも簡単に視聴・投稿でき

[†] 電気通信大学大学院情報システム学研究科
Graduate School of Information Systems, The University of Electro-Communication
^{††} 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology

るようになった。また、ニコニコ動画等の動画共有サイトを国民の7割以上が利用しているという調査結果[1]も発表されており、ニコニコ動画等の動画共有サイトは一般の人々に身近な存在となってきている。しかし、大量の情報を誰もが大量に取得できるという状況が生じたために、かえって欲しい情報を発見できないという情報過多 (information overload)の問題が指摘されている[14]

ニコニコ動画では、再生数 (視聴したユーザの数) やマイリスト数 (お気に入りに登録したユーザの数) によって動画の重み付けを行っており、この重み付けを利用したランキング機能によって、動画の発見が支援されている。しかし、ランキングに掲載されるためには、多くのユーザに視聴され万人受けされる必要があり、小規模・中規模のユーザ達が楽しんで視聴している動画の抽出ができないという問題がある。また、ランキングに掲載されたことによって、掲載された動画コンテンツに対する支持が一層強固なものになってしまうバンドワゴン効果の問題[17]もある。

本研究では、多くの人には知られていないが、視聴した際に多くの人の興味・関心が湧く動画コンテンツを発見したいと考える。そこで、ノベルティ[2]のあるコンテンツ (ユーザが知らないかつ好みのコンテンツ)の意味に則り、「社会的には知られていないが、より多くの人が興味・関心を持つコンテンツ」をソーシャルノベルティ (Social-Novelty)のあるコンテンツであると定義し、このような動画の発見を目指す。ソーシャルノベルティのある動画は、個々のユーザに則した推薦ではなく、社会的に意外性・新規性のある動画である。そのためそのような動画の発見はユーザの満足度の向上に繋がり、ひいてはシステムの活性化に繋がると考えている。

本論文の構成を以下に示す。2章では、従来の推薦システムとニコニコ動画に関する先行研究について述べる。3章では、ソーシャルノベルティのある動画を発見するために動画の3特性に注目し、ソーシャルノベルティのある動画を発見する方法を提案する。4章では、ソーシャルノベルティのある動画をもっと評価されるべきタグが付与された動画であると仮定し、「もっと評価されるべき」タグについて分析を行う。5章では、4章の分析結果を踏まえ「もっと評価されるべき」タグが付与された動画のコメントの特徴語の分析を行う。6章では、ソーシャルノベルティのある動画を自動分類するために、機械学習の精度の分析を行う。7章においてまとめ、8章で今後の課題を述べる。

2. 関連研究

2.1 従来の推薦システムに関する研究

推薦を行うにあたり、推薦システムは大きく分けて個人化と非個人化の2つに分けられることが知られている[14]。

個人化の推薦システムの代表的なものとして、協調フィルタリング[3][4][14]やコン

コンテンツベースフィルタリング[4][5][14]が知られている。協調フィルタリングではコンテンツの選択・評価履歴から、当該ユーザと似ているユーザ群を発見し、コンテンツの選択履歴を基にコンテンツを推薦する。コンテンツベースフィルタリングでは、コンテンツからテキスト等の特徴量を抽出することで、コンテンツ間の類似度を計算し、当該ユーザのコンテンツの選択履歴を基に類似しているコンテンツを推薦する。これらの推薦手法では正確さを推薦手法の評価指標として用いている。

しかし、正確さを評価指標とするだけでは不十分であることが述べられている[2][6]。なぜなら個人の嗜好に合わせた推薦が行われても、ユーザが知っているコンテンツであれば、それが有用な推薦であるとはいえないからである。正確さ以外の評価指標を用いた推薦システムとして、ノベルティ[2][7]やセレンディピティ[8]を用いた推薦手法がある。

一方、非個人化の推薦とは、すべてのユーザに全く同じ推薦をすることを指す[14]。例えば、ニコニコ動画においてはランキングが非個人化の推薦に該当する。しかしランキングでは、今流行っている動画、多くのユーザが知っているカテゴリの動画、人気のある投稿者が投稿した動画等がランキングに掲載されることが多いと考えられるため、ノベルティは小さいと考える。我々は「社会的には知られていないが、より多くの人が興味・関心を持つコンテンツ」を発見したいため、非個人化推薦の立場をとりながら、ノベルティの大きい動画の発見を目指している。

本研究と既存の推薦手法の位置づけをまとめると図1のようになる。個人化の推薦

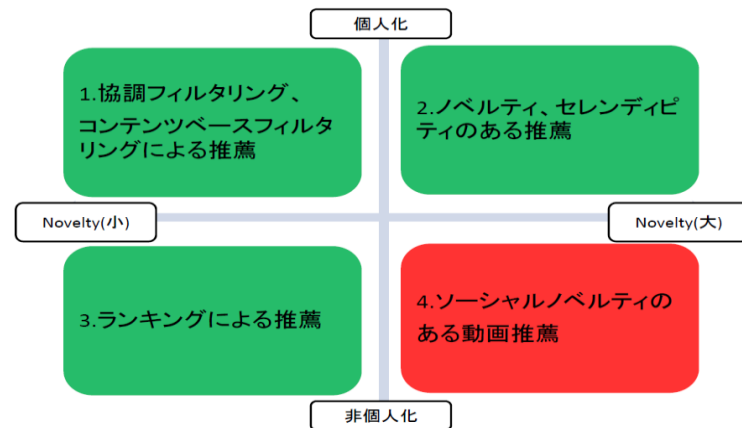


図1 本研究と既存の推薦手法の位置づけ

システムとして協調フィルタリングやコンテンツベースフィルタリングがあり、それらをベースとしてノベルティやセレンディピティに注目した研究が展開されている。これらは個人化の推薦であるため、社会的な立場から推薦を行いたい我々とは立場が異なる。非個人化の推薦システムとしてはランキングが挙げられるが、ノベルティは小さいと考える。そこで、非個人化の推薦でありながらノベルティの大きなソーシャルノベルティのある動画を発見することを本研究の目的とする。

2.2 ニコニコ動画に関する研究

ニコニコ動画に関する研究として、タグに注目することで、タグ共起ネットワークの特徴を抽出し可視化した研究[10]や、「初音ミク」というタグに注目して協調的な創造活動を明らかにした研究[16]、また視聴者の共感度を抽出・定量化するためにタグの頻度や付き方の分析を行った研究[11]がある。更に、投稿されたコメントの時系列に注目することで、動画の盛り上がっている部分を要約する研究[15]もある。これらの研究はニコニコ動画のタグやコメントに注目することで、動画の特徴やユーザの振る舞いの特徴、あるいはユーザにとって有意義なアウトプットが抽出できることを示している。本稿でも先行研究と同様にコメントやタグに注目して分析を行うが、目指すアウトプットがソーシャルノベルティのある動画であることが先行研究とは異なる。

3. ソーシャルノベルティのある動画の発見手法

3.1 ソーシャルノベルティのある動画

動画共有サイトに投稿されている動画は様々であり、玉石混交な動画が投稿されている。投稿される動画を、「再生数の大きさ」と「興味・関心を持った人数」に基づき4つのタイプに分類する(図2)。第一象限の動画は、再生数が大きく、かつ興味・関心を持つ人数が多い動画である。興味・関心を持つ人数が多いことから、視聴価値の高い動画と考える。このような動画は、いわゆる玉石混交の「玉」の動画である。第二象限の動画は、再生数は大きい、興味・関心を持つ人数が少ない動画である。興味・関心を持つ人が少ないことから、視聴価値が低い動画と考える。視聴価値が低いにもかかわらず再生数が増えるのは、工作行為によって不正に再生数が引き上げられているものと考えられる。よって、このような動画は石を玉に見せかけた「紛い物」といえる。第三象限の動画は、再生数が小さく、興味・関心を持つ人数が少ない動画である。興味・関心を持つ人数が少ないことから、視聴価値は低いと考える。このような動画は、玉石混交の「石」の動画である。第四象限の動画は、再生数が少ないが興味・関心を持つ人数が多い動画である。興味・関心を持つ人数が多いことから視聴価値が高いと考える。このような動画は、一見石に見える玉であり、掘り出し物といえる。掘り出し物は、再生数が少ないことから社会的に知られていないと考えられ、かつ興味・関心を持つ人数は多いことからソーシャルノベルティのある動画といえる。

「玉」の動画は、視聴価値は高いがランキングによって推薦されているため本稿では対象としない。「紛い物」や「石」の動画は、視聴価値が低いため、そもそも推薦する必要はなく、対象外である。我々は、ソーシャルノベルティのある動画を推薦対象とし、次節でその手法を提案する。

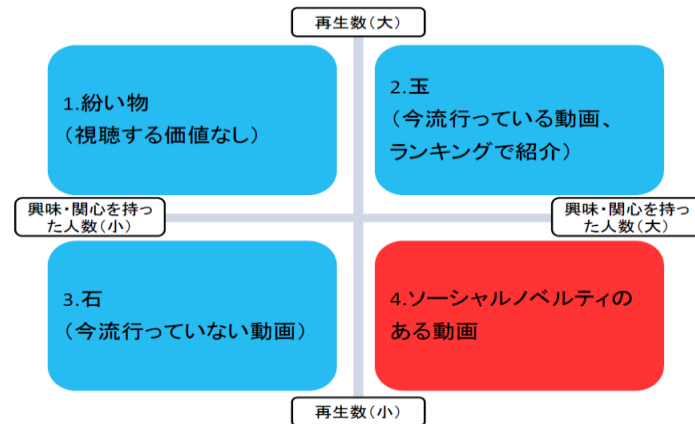


図2 ソーシャルノベルティのある動画の位置づけ

3.2 提案手法

ニコニコ動画には動画に対して感想等を自由に書きこむことができる「コメント機能」と、動画の性質や特徴等を表す言葉を自由に登録・編集することができる「タグ機能」がある。また、動画が投稿されてから現在までの累計再生数、コメント数（コメントを残したユーザの数）、マイリスト数も記録されている。我々はこれらのコメントやタグ、再生数等に注目することでソーシャルノベルティのある動画が発見できるのではないかと考えた。そこで、動画の特徴づける特性として、以下の3つの特性に注目する。

1. コメント特性：コメントに注目

コメント機能はニコニコ動画の特徴的な機能であり、ユーザー一人一人の生の声を表している。従って、ソーシャルノベルティのある動画にはそれを表す特徴的な表現があると考えられる。そこで、コメントの特徴語やコメントの投稿頻度を、動画の特徴付ける特性として用いることで、ソーシャルノベルティのある動画が発見できると考える。

2. 動画内容特性：タグに注目

タグは、動画の内容を端的に表している。ソーシャルノベルティのある動画には、それと関連するタグが付与されていると考える。よって、タグの有無や共起パターンを用いることで、ソーシャルノベルティのある動画を発見できると考える。

3. 視聴行動特性：再生数、コメント数、マイリスト数に注目

再生数、コメント数、マイリスト数は、ユーザが視聴行動をすることで増減する。TV番組における視聴率のように、ソーシャルノベルティのある動画には、それを表す特徴的な視聴行動特性があると考えられる。よって、視聴行動特性からソーシャルノベルティのある動画を発見できると考える。

これらの3つの特性を組合わせて用いることで、ソーシャルノベルティのある動画が発見する方法を提案する（図3）。本研究ではソーシャルノベルティのある動画を、機械学習を用いて自動的に分類することを目指す。

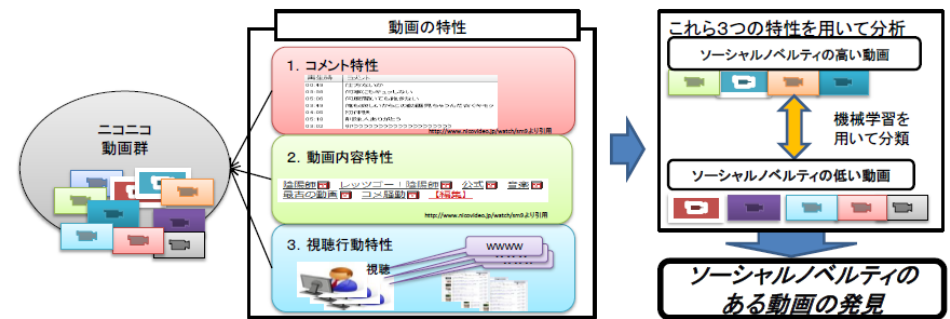


図3 提案手法の概要図

4. ソーシャルノベルティのある動画の分析

ソーシャルノベルティのある動画発見のために活用できるタグは「もっと評価されるべき」タグであると考えられる。なぜなら「もっと評価されるべき」タグは、「動画が良質で面白いのにも関わらず再生数等の評価が低い、あるいは存在自体認知されていないと感じられた場合に、視聴者側から自然発生的に貼りつけられるタグ」[9]であるといわれているからである。

以下の分析ではまず「もっと評価されるべき」タグが付与された動画がソーシャルノベルティのある動画の正解データとしてよいかを判断するために、「もっと評価され

るべき」タグについて分析を行う。

4.1 分析に用いるデータ

動画データを収集するために php を用いてプログラムを作成し、ニコニコ動画 API を用いてクローリングを行った。クローリングを行った期間は、2011年5月2日から2011年5月6日で、クローリングを行った動画データは2009年9月1日から2009年9月5日に投稿された動画16,770件である。収集したデータの内容は、タイトル、動画時間、投稿日時、再生数、コメント数、マイリスト数、タグ、コメント内容である。これらのデータを分析に用いる。

ソーシャルノベルティのある動画は、ランキングに掲載されていない動画を対象としている。そのため、ランキングに掲載されている動画に関しては、分析対象から除外する必要がある。2010年1月1日～2010年6月30日におけるニコニコ動画デイリーランキング総合300位の再生数・コメント数・マイリスト数の値を調査した結果、それぞれの値の平均値は314,616.2件、23,375.6件、12,334.3件であった。よって、クローリングを行った動画データの中で再生数・コメント数・マイリスト数がそれぞれ30万、2万、1万のいずれかを越えている動画に関しては以降の分析対象から除外する。

4.2 分析方法／結果

まず、「もっと評価されるべき」タグがニコニコ動画内でユーザに活用されていることを確認するために、タグが使用されている回数を算出した(表1)。この中で、「ゲーム」、「実況プレイ動画」、「音楽」、「歌ってみた」、「東方」、「アニメ」、「エンターテイメント」、「アイドルマスター」タグは、ニコニコ動画公式のカテゴリ分けするために用意されているタグである。また、「実況プレイ」タグは「ゲーム」タグと共に用いられることが非常に多い。ソーシャルノベルティのある動画の発見のために、これらの個人の嗜好に依存するタグを利用するのは得策ではない。これらのタグを除外すると、「もっと評価されるべき」タグはカテゴリに依存しないタグの中で最も活用されているタグであることが分かる。

次に、ユーザの興味・関心の度合いが高いタグを発見するため、使用回数が100回以上のタグについて平均マイリスト数を算出した(表2)。マイリスト数に注目した理由は、マイリスト登録が興味・関心が湧いた動画に対してユーザが行う行為の一つであると考えたからである。表2において、「演奏してみた」、「作業用BGM」、「Vocaloid」、「東方」、「歌ってみた」、「アニメ」、「音楽」タグは先程同様ニコニコ動画公式のカテゴリタグであり、「スネオが自慢話をするとき流れている曲」タグは、「音楽」タグと共に用いられることが非常に多い。表1の結果同様、「もっと評価されるべき」タグはカテゴリに依存しないタグの中でマイリスト数の値が高いタグであることが分かる。

次に、「もっと評価されるべき」タグが付いている動画は、付いていない動画に比べると平均マイリスト数が高くなっていることを確認するため、タグの有無とランキン

グの掲載の有無に関して平均マイリスト数と動画数を算出した(表3)。なお、ニコニコ動画で多く使われている「音楽」「ゲーム」タグを比較対象として掲載している。この結果から、「もっと評価されるべき」タグは付いていない動画に比べてマイリスト数が多く、ランキングにも入ってこないことが分かる。

また、「もっと評価されるべき」タグそのものが、ニコニコ動画公式のカテゴリタグに依存していないかを調査した。「もっと評価されるべき」タグが付与されている動画454件中、20回以上用いられたタグの共起頻度を算出した(図4)。この結果から、「もっと評価されるべき」タグは特定のカテゴリに依存することなくユーザに利用されていることが分かる。

更に、先行研究において、ニコニコ動画のランキングに掲載されている動画を8日分収集し、タグ毎に度数分布の上位10位までに入った回数を日毎に算出した研究がある[10]。この研究の中でも「もっと評価されるべき」というタグはマイリスト登録される傾向が高いことが述べられている。また、ニコニコ動画のタグの付き方を解析した研究[11]において、「もっと評価されるべき」タグは投稿者によってロックされている割合がとて低いことが述べられている。(ニコニコ動画のタグ機能は、投稿者に限りタグを動画にロックすることができる。)

表1 タグの使用回数上位10

タグ	使用回数
ゲーム	8,202
実況プレイ動画	2,939
音楽	1,345
歌ってみた	749
東方	499
アニメ	459
もっと評価されるべき	454
エンターテイメント	409
アイドルマスター	360
実況プレイ	340

表2 タグのマイリスト数の平均値上位 10

タグ	マイリスト数の平均値
演奏してみた	530.0
作業用 BGM	249.9
初音ミク	213.6
Vocaloid	212.2
東方	161.0
もっと評価されるべき	153.8
歌ってみた	148.5
アニメ	140.4
音楽	139.7
スネ夫が自慢話をするとき 流れている曲	134.3

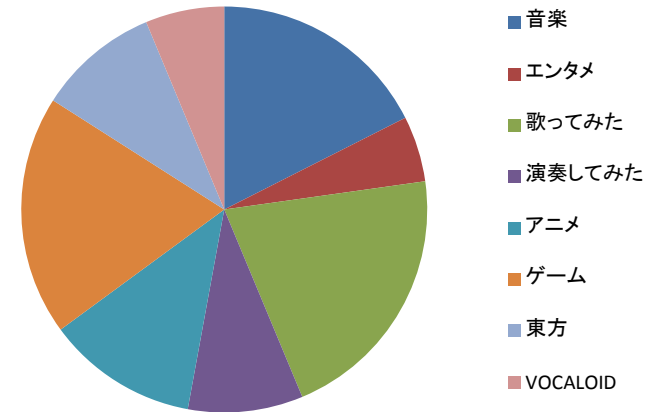


図4 「もっと評価されるべき」タグとの共起関係

表3 タグの平均マイリスト数と動画数の比較
(表の値は平均マイリスト数, 括弧は動画数)

タグ	ランキングに入っていない動画 75(15,636)		ランキングに入っている動画 12,436.7(34)	
	当該タグが 付いていない	当該タグが 付いている	当該タグが 付いていない	当該タグが 付いている
もっと評価 されるべき	72.6(15,182)	153.8(454)	12,436.7(34)	該当動画なし
音楽	42.8(14,291)	130.7(1,345)	12539.0(31)	11,380.3(3)
ゲーム	121.5(8,202)	23.7(7,434)	19,038.1(20)	2,942.0(14)

4.3 分析結果のまとめ

「もっと評価されるべき」タグがソーシャルノベルティのある動画の正解データとして用いることが適切かどうかに関して、本分析と先行研究をまとめると「もっと評価されるべき」タグは以下の性質を持っている。

- i. マイリスト数が多い。つまり、興味・関心を持つユーザーが多いと考えられる。
- ii. ランキングに掲載されない。すなわち、再生数は低い。タグが付かない理由は、ランキングに掲載されるとユーザーが「もっと評価されるべき」タグを削除する自浄作用が働いているからと考えられる。
- iii. 多くのユーザーに利用されている。つまり、ソーシャルノベルティのある動画をユーザーが発見した時にこのタグを付けることができる場合が多いということが考えられる。
- iv. 特定のカテゴリに依存していない。そのため、社会的に知られていない動画をカテゴリに縛られず発見したい本研究の目的と合致する。
- v. 投稿者によってロックされている割合が低い。つまり、ユーザーが自由に登録・削除しているタグであるため、動画の状態を正しく表せていると考えられる。投稿者にロックされていると「もっと評価されるべき」タグが付いた動画が評価され、ランキングに掲載された場合にこのタグを削除できない（自浄作用が働かない）という状況に陥る。

以上の結果を踏まえ、「もっと評価されるべき」タグが付いた動画は、興味・関心を

もったユーザの割合が多い一方で、再生数は低いため、他の多くのユーザは知らないと言える。また、使用頻度が高いため、ニコニコ動画において代表的なタグであると考えられ、どのカテゴリにも使われているため、汎用性もある。更に投稿者からロックをされていないため、ユーザは「もっと評価されるべき」タグを追加したいとき、削除したいときに自由に追加、削除することができる。これらの分析結果を踏まえて「もっと評価されるべき」タグをソーシャルノベルティの動画の分類に利用できると判断した。

5. コメント特性を用いたコメントの特徴語の分析

コメント特性を用いてソーシャルノベルティのある動画を発見するためには、それらの動画のコメントの分析を行う必要がある。機械学習を行う前の事前調査として、「もっと評価されるべき」タグが付与された動画とされていない動画ではコメントの特徴にどのような違いがあるのかを分析した。形態素解析プログラムとして Mecab を用いて、動画の特性を強く表しているであろう名詞と形容詞について単語の抽出を行う。

5.1 分析方法／結果

「もっと評価されるべき」タグが付与されている動画の件数、合計コメント数、平均コメント数はそれぞれ 454 件、109,839 件、241.9 件、それ以外の動画の件数、合計コメント数、平均コメント数はそれぞれ 15,670 件、4,221,130 件、261.7 件だった。これらの動画データのコメントに形態素解析を行い、頻度の高かったそれぞれの上位 100 語に関して、それぞれの動画にのみ現れた語を以下の表 4 に示す。なお、情報分析システム WISDOM[12][13]を使用して、抽出された言葉をポジティブとネガティブにも分類した。

「もっと評価されるべき」タグが付与されている動画では"評価"や"素晴らしい", "かっこいい"等のポジティブな言葉が、それ以外の動画では"hidden"や"こいつ"等のネガティブな言葉が多く抽出された。コメントは動画に対するユーザの率直な感想であると考えられるので、ポジティブな言葉が多く抽出された「もっと評価されるべき」タグが付与されている動画は、ソーシャルノベルティのある動画としてユーザに推薦された場合に、表 4 で示したようなポジティブな感想を抱くことが予想される。それ以外の動画は、推薦された場合に視聴した際にネガティブな感想を抱くことが予想される。

6. 機械学習の精度分析

6.1 分析方法

「もっと評価されるべき」タグが付与された動画データを正解データとして機械学

表 4 コメントの特徴語抽出結果
(括弧は頻度、下線はポジティブな言葉、斜体はネガティブな言葉)

順位	「もっと評価されるべき」タグが 付与されている動画	それ以外の動画
1	<u>評価</u> (361)	<i>hidden</i> (881)
2	<u>素晴らしい</u> (133)	子(265)
3	<u>かっこよい</u> (127)	<i>ゲーム</i> (242)
4	<u>H</u> (112)	日本(223)
5	<u>すばらしい</u> (109)	<i>こいつ</i> (215)
6	<u>最高</u> (103)	レベル(211)
7	<u>夢</u> (96)	目(205)
8	嫁(85)	<i>きつい</i> (200)
9	<u>GJ</u> (84)	<i>弱い</i> (199)
10	<u>綺麗</u> (84)	相手(192)
11	涙(83)	キャラ(192)
12	<u>素敵</u> (82)	<i>ダメ</i> (192)
13	<u>イイ</u> (80)	<i>低い</i> (191)
14	<u>歌詞</u> (77)	<i>難しい</i> (189)
15	<u>HIT</u> (72)	頭(181)
16	<u>かわゆい</u> (70)	<i>なに</i> (173)
17	<u>大好き</u> (68)	<i>熱い</i> (170)
18	<u>君</u> (67)	<i>だめ</i> (168)
19	鳥肌(65)	<i>酷い</i> (167)
20	<u>優しい</u> (65)	<i>遅い</i> (164)

習の精度を測る。機械学習のアルゴリズムとして、教師ありによる機械学習サポートベクターマシンを、実行には WEKA を用いた。5章で抽出したコメントの特徴語を機械学習に利用し、更に視聴行動特性の再生数・コメント数・マイリスト数の中でどの視聴行動特性を用いるのが精度が上がるのかを検証した。なお、10分割交差検証法により精度を算出をしている。また、機械学習に用いたデータは、「もっと評価されるべき」「音楽」タグが付与されている動画 50 件、「音楽」タグが付与されている動画 200 件、計 250 件である。

6.2 分析結果

分析結果は表5のようになり、コメント内容とマイリスト数を組み合わせた場合が最も精度が高くなった。つまり、ソーシャルノベルティのある動画を分類するためには、コメント内容とマイリスト数を用いることが最も有効であることが分かる。また、視聴行動特性においてはマイリスト数の値が最もソーシャルノベルティを表しているということも分かる。

表5 機械学習の精度比較結果

機械学習に用いたデータ	適合率
コメント内容のみ (コメントの特徴語を活用)	78.8
マイリスト数のみ	65.3
コメント内容, 再生数	79.7
コメント内容, コメント数	79.4
コメント内容, マイリスト数	80.5
コメント内容, (コメント数÷再生数)	79.7
コメント内容, (マイリスト数÷再生数)	79.1
コメント内容, (マイリスト数÷コメント数)	79.1

6.3 考察

視聴行動特性のみを用いるのでは、精度が低い結果となった。本分析では投稿日が異なる動画であっても、再生数、コメント数、マイリスト数が同じ値であれば同列に扱っている。したがって、投稿日が最近の動画の方が、再生数、コメント数、マイリスト数の値が同じだった場合、価値があると考えられる。よって、投稿日によって再生数、コメント数、マイリスト数にバイアスをかけるべきであると考えられる。

各精度間に統計的に有意な差があるかを検証するために統計的検定を行ったが、有意な差はなかった。これは、カテゴリ毎に視聴行動特性に偏りがあることが関係しているからと考えられる。したがって、今後は動画内容特性であるタグを分析内容に加味することで他のカテゴリにも注目し分析を行う。また、データ数を増やして提案手法を適用させる必要がある。

また、分類した動画の評価実験を行っていないので、本提案の有用性は不明である。したがって今後、分類された動画の評価する実験を行わなければならないと考えている。

7. まとめ

本研究ではソーシャルノベルティのある動画を発見するために、動画の3特性に注目した。これらの特性に基づき、「もっと評価されるべき」タグに注目し、「もっと評価されるべき」タグがソーシャルノベルティのある動画の正解データとして利用できることを確認した。また、「もっと評価されるべき」タグが付与されている動画のコメントの特徴語を抽出し、その特徴語はポジティブな言葉が多いことを確認した。そしてソーシャルノベルティである動画を分類する精度が最も上がったのは、コメント内容とマイリスト数を組み合わせた場合であることを分析により明らかにした。

8. 今後の課題

本稿では動画内容の特性について分析を行っていない。今後は「もっと評価されるべき」タグの付いている動画において、タグの頻度や共起関係をその他の動画と比較し、機械学習に用いることを検討する。

また、発見したソーシャルノベルティのある動画をユーザに視聴してもらい、マイリスト登録を行いたい人の割合を調査する評価実験を行う予定である。

参考文献

- 1) Recording Industry Association of Japan 「動画サイトの利用実態調査検討委員会」報告書 <http://www.riaj.or.jp/release/2011/pdf/20110808report.pdf>
- 2) Herlocker, J., Konstan, J. and Riedl, J.: Evaluating collaborative filtering recommender systems, ACM Trans. Information Systems, Vol.22, No.1, pp. 5-23(2004)
- 3) Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering, IEEE Internet Computing, pp. 76-80(2003)
- 4) 土方嘉徳: 嗜好抽出・情報推薦の基礎理論1. 嗜好抽出と情報推薦技術, IPSJ Magazine, Vol.48, No.9, pp. 957-965(2007)
- 5) Blanco-Fernandez, Y., Pazos-arias, J., Gil-Solla, A., Ramos-Cabrer, M., Lopez-Nores, M.: Providing entertainment by content-based filtering and semantic reasoning in intelligent recommender systems, IEEE Transactions on Consumer Electronics, pp. 727-735(2008)
- 6) McNeel, S.M., Riedl, J. and Konstan, J.: Being accurate is not Enough: How Accuracy Metrics have hurt Recommender Systems, A Work-In-Progress paper in the Extended Abstracts of the 2006 ACM Conf. on Human Factor in Computing Systems (ACM SIGCHI), pp. 1097-1101 (2006)
- 7) 小川祐樹, 諏訪博彦, 山本仁志, 岡田勇, 太田敏澄: 動的なトピック分類に基づく Novelty を考慮した推薦アルゴリズムの提案, 情報処理学会論文誌, Vol.50, No.6, pp. 1636-1648 (2009)
- 8) 秋山高行, 小原清弘, 谷崎正明: ユーザの選択履歴に依存しない指標を利用した serendipity のある推薦方式の提案と評価, 情報科学技術フォーラム講演論文集, Vol.9, No.4, pp.157-164 (2010)

- 9) . ニコニコ大百科「もっと評価されるべき」(2011年9月1日参照)
<http://dic.nicovideo.jp/a/%E3%82%82%E3%81%A3%E3%81%A8%E8%A9%95%E4%BE%A1%E3%81%95%E3%82%8C%E3%82%8B%E3%81%B9%E3%81%8D>
- 10) 伊藤聖修,鈴木育男,山本雅人,古川正志: ニコニコ動画におけるタグ共起ネットワークの特徴抽出,人工知能学会研究会資料,知識ベースシステム研究会 vol.80, pp.13-18 (2008)
- 11) 村上直至,伊東栄典: 動画投稿サイトで付与された動画タグの階層化,情報処理学会研究報告, バイオ情報学 Vol.23, No.17, pp.1-6 (2010)
- 12) 情報分析システム WISDOM, <http://wisdom-nict.jp/>
- 13) 河原大輔,黒橋禎夫,乾健太郎: 主要・対立表現の俯瞰的把握: ウェブの情報信頼性分析に向けて,情報処理学会研究報告.自然言語処理学会研究報告, Vol.67, pp.49-54 (2008)
- 14) 神鷲敏弘: 推薦システムのアルゴリズム(1),人工知能学会論文誌,Vol.22, No.6, pp.826-837(2007)
- 15) 青木秀憲,宮下芳明: ニコニコ動画における映像要約とサビ検出の試み,情報処理学会研究報告, ヒューマンコンピュータインタラクション研究会報告, Vol.50, pp.37-42(2008)
- 16) 濱崎雅弘,武田英明,西村拓一: 動画共有サイトにおける大規模な協調的創造活動の創発のネットワーク分析, 人工知能学会論文誌, Vol.25, No.1, pp.157-167 (2010)
- 17) 小川祐樹,諏訪博彦,太田敏澄: 主体性の拡張を支援する推薦システムに関する研究, 日本社会情報学会第 24 回全国大会論文集,pp.372-375