

推薦論文

コンテンツのユーザに与える感情を考慮した コンテンツ推薦手法の検討

藤田 俊^{1,a)} 鷹野 孝典²

受付日 2019年7月16日, 採録日 2020年3月13日

概要: マイクロブログやレビューサイトなどの情報サービス上の膨大な情報からユーザの共感や興味を喚起する情報を提供することは重要である。情報サービス上の発言には、感想や評価などのようなユーザがコンテンツに対して抱いた感情を含む発言が多く存在する。このような感情を考慮することで共感や興味を抱くコンテンツの推薦が可能である。しかし、協調フィルタリングのように類似したユーザから推定される感情を考慮するだけでは、推薦対象となるコンテンツがユーザの感情に与える影響は十分に考慮されていないため、コンテンツが与える感情との関係を考慮する仕組みが必要である。本研究では、コンテンツのユーザに与える感情を考慮したコンテンツ推薦手法を提案する。提案手法は、コンテンツに対してユーザが抱く感情およびコンテンツが与える感情を、ユーザの発言履歴から感情判定モデルにより8種の感情値として算出する。推薦処理では協調フィルタリングを適用することにより、ユーザの感情間の類似性に基づいて、未知のコンテンツに抱く感情を算出し、コンテンツが与える感情との関係性を考慮したコンテンツ推薦を行う。ユーザが抱く感情に類似した感情を与えるコンテンツを提示することで、そのコンテンツに対するユーザの興味を高める効果が期待できる。本研究では、マイクロブログの1つであるTwitterを対象とした実験により、提案手法の実現可能性と有効性を検証する。

キーワード: 感情分析, 情報推薦, コンテンツ, マイクロブログ, Twitter

A Method for Content Recommendation Using User's Emotions Derived from Contents

SHUN FUJITA^{1,a)} KOSUKE TAKANO²

Received: July 16, 2019, Accepted: March 13, 2020

Abstract: It is important to provide information that evokes users' empathy and interest from the vast amount of information on information services such as microblog and review sites. It would be possible to recommend content that makes a user have empathy and interest to it by considering such emotions, however, it is insufficient to consider only emotions predicted from similar users like collaborative filtering, since it does not consider the influence that the content to be recommended will give on user's emotion. In this study, we propose a method for content recommendation using user's emotions derived from contents. The proposed method calculates the emotion held by a user for a content and the emotion given by the content as eight types of emotion values from a history of user's utterances. In the recommendation process, collaborative filtering is applied to calculate emotions in unknown content based on the similarity between emotions of users, and suitable contents are recommended based on the relationship with emotions given by content. In this study, we verify the feasibility and effectiveness of the proposed method through experiments using Twitter, which is one of microblog services available on the Web.

Keywords: emotion analysis, information recommendation, content, microblog, Twitter

¹ 神奈川工科大学大学院博士前期課程情報工学専攻
Course of Information and Computer Sciences, Graduate
School of Engineering, Kanagawa Institute of Technology,
Hon-Atsugi, Kanagawa 243-0292, Japan

² 神奈川工科大学情報学部情報工学科
Department of Information and Computer Sciences, Faculty
of Information Technology, Kanagawa Institute of Technol-
ogy, Hon-Atsugi, Kanagawa 243-0292, Japan
a) syun070822@gmail.com

1. はじめに

オンラインショッピングサイトや音楽・映画の配信サイトでは、膨大な商品、楽曲、動画に関するコンテンツ情報が存在するため、ユーザが求めるコンテンツ情報をより効率的に提供するために情報を推薦することが求められる。情報推薦の分野では、協調フィルタリングのように、ユーザの嗜好情報に基づいて複数のユーザ間の類似性を考慮することにより、ユーザが興味を持つと予測されるコンテンツ情報を抽出し、推薦する手法が有効であると考えられる。

一方、Twitter, Facebook, Amazon などのマイクロブログやレビューサイトなどの情報サービスを利用することで、ユーザは情報の発信や他ユーザの発言を閲覧することができる。このような情報サービス上の発言には、感想や評価などのようなユーザが商品、楽曲、映画などのコンテンツに対して抱いた感情が含まれる発言が多く存在する。ユーザ感情を考慮した情報推薦に関する研究としては、ユーザが抱く複数の感情の類似度を考慮したユーザ推薦手法 [1], [2] がある。また、ユーザの感情だけでなく、コンテンツが人に与える感情を考慮した推薦手法の研究 [4], [5] も存在する。以上の先行研究をふまえて、コンテンツ推薦においては、対象ユーザが推薦コンテンツに対して抱く複数の感情とその推薦コンテンツが内容に基づいてあらゆるユーザに与える複数の感情の関係性を考慮した推薦をすることで、推薦されたコンテンツが与えると考えられる感情に対する共感や興味を持たせることができるのではないかと考えられる。

この観点において、既存のコンテンツ推薦手法には、下記の2つの問題点がある。

[課題 1] 感情を考慮した情報推薦としては、ポジティブとネガティブの2極値を参照して行う手法が提案されている。しかし、ポジティブとネガティブの2極値では、複数の感情を考慮する場合には不十分である。たとえば、同じポジティブな感情でも、あるコンテンツに対しては「面白い」といった感情を抱いているが、別のコンテンツに対しては「驚き」といった異なる感情を抱く場合もあれば、1つのコンテンツに対して「面白い」、「興奮」といった感情を同時に抱くような場合もあり、このようなユーザ嗜好の傾向を考慮した情報推薦に十分に対応することができない。そのため、まず、ユーザが特定のコンテンツに対して抱く「面白い」、「驚き」といった複数の種類の感情を類推し、コンテンツ推薦のための指標として考慮する必要がある。

[課題 2] 協調フィルタリングでは類似した感情を持つユーザを抽出し、未知のコンテンツに対して抱くと予想される感情を類推することは可能であるが、推薦対象となるコンテンツがユーザの感情と一致するコンテンツであるかどうかを判定しない。このため、ユーザが抱く感情に訴えかけ、共感を持たせるコンテンツ推薦を行うためには、ユー

ザがコンテンツに抱く感情とそのコンテンツがユーザに与える感情との関係を考慮する仕組みが必要となる。ここで「ユーザがコンテンツについて抱く感情」とは、対象ユーザが特定のコンテンツに対して抱く感情である。このとき、ユーザにとって未知のコンテンツも対象とするので、必ずしもコンテンツがユーザに対して与える情報に基づいて抽出される感情情報とは限らない。一方、「コンテンツが与える感情」は、特定のコンテンツが与える情報に対して、不特定多数のユーザが持つであろう感情情報を表している。

本研究では、コンテンツのユーザに与える感情を考慮したコンテンツ推薦手法を提案する。提案手法では、上記課題 1 の複数の感情を考慮するために、マイクロブログにおける発言履歴からコンテンツに対して抱くユーザ感情を8軸の感情表現モデルに基づいて抽出し、さらにコンテンツがユーザに与えると期待される平均的な感情を算出することで、ユーザおよびコンテンツの感情に関するメタデータ・データベースの構築を行う。推薦処理では協調フィルタリングを適用することにより、ユーザの感情間の類似性に基づいて、未知のコンテンツに抱く感情の算出に加え、上記の課題 2 を考慮した、コンテンツが与える感情との関係性を考慮したコンテンツ推薦を行う。ユーザが抱く感情に類似した感情を与えるコンテンツを提示することで、従来手法ではユーザの感情のみを考慮した情報推薦は可能だったが、実現できなかった推薦対象となるコンテンツがユーザの感情と一致するコンテンツであるかどうかを考慮した推薦が可能となる。たとえば、推薦対象となっている未知のコンテンツに対してユーザが「哀しい」といった感情を抱くと判定さ、同時にそのコンテンツが「哀しい」、「感動」といった感情をあらゆるユーザに与えると判定される場合において、推薦対象コンテンツに対するユーザの興味を高める効果が期待できる。

本研究では、Twitter を対象とした実験により、提案手法の実現可能性と有効性を検証する。

2. 関連研究

2.1 感情値を利用した情報推薦手法

抽出した複数の感情を用いて共通の話題を持つユーザを推薦する手法も研究されている。熊本らの手法 [1], [2] では、文章を構成する各単語の感情の強さをあらかじめ感情軸ごとに定量化した感情語辞書を基に、ユーザのツイートから感情値のベクトルを生成し、共通の話題を持つユーザとコサイン類似度により類似する感情を持つユーザをフォロワー候補として推薦している。加藤ら [3] はユーザごとのすべての単語に対して喜・怒・哀の3種類の感情に分類し、マッピング化しつつコサイン類似度を用いて類似する

本論文の内容は 2018 年 9 月の第 176 回マルチメディア通信と分散処理研究会発表会にて報告され、同研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

単語を持つユーザを推薦する手法を提案するとともに単語ごとの感情値の傾向を確認している。

また、類似する感情を考慮したコンテンツ推薦についても研究されている。村石らはネガティブな感情を抱いているユーザに対して対となる感情を誘発する行動を推薦する手法を提案している [4]。佐藤らの手法は入力文章から8種類の感情を推定し、コサイン類似度を用いて類似する感情ベクトルを持つことわざを推薦する [5]。平良らは得たい感情を入力してもらい、小説の認知評価および感情パターンの蓄積、入力した短文からの感情推定、データのマッチングの3段階によりその感情を得られると予測される小説を推薦する手法を提案し、仮想の質問セットと仮想ユーザとのマッチングによるシミュレーションを行っている [6]。

2.2 感情の抽出手法

テキスト情報から複数の感情の分類や感情値を基に特徴を抽出する手法が研究されている。杉本らの手法は感情語を8つのカテゴリに分類した感情語辞書を作成し、口コミに含まれる単語とのマッチングを行うことで口コミに含まれる感情を抽出する [7]。若宮らはWeb ページ検索において、検索語により得られたWeb ページから重要度の高い単語を補完語として抽出し、補完語を利用した再検索により、感情の多様性を考慮したWeb ページの検索を実現した [8]。橋本らは従来のポジネガ判定に10軸の感情を対象とした感情分析を併用することで時系列変化による評判の変化を抽出する手法を提案している [9]。Colnericらは3つの感情データセットを作成し、単語および文字ベースの畳み込みニューラルネットワークの性能を単語のバックワードおよび潜在意味索引付けモデルの性能と比較した結果、特に文字ベースの畳み込みニューラルネットワークが語彙や潜在意味索引付けモデルを改善できることを確認した [10]。

2.3 協調フィルタリングと組み合わせた推薦手法

協調フィルタリングに何らかの手法を組み合わせることで性能を向上させる推薦手法も研究されている。

服部らは既存の協調フィルタリングに価値観モデルを適用した協調フィルタリングを組み合わせることで、予測誤差およびユーザ・アイテムの被覆率が改善されたことを確認している [11]。西村らは協調フィルタリングにトピックモデルを融合させた Collaborative Topic Regression (CTR) にユーザに関するテキスト情報を用いた評価がされていない点に着目し、CTR にアイテムに関するテキスト情報を用いた場合とユーザに関するテキスト情報を用いた場合の結果を比較することで、アイテムの場合は被覆率および多様性が高く、ユーザの場合は再現率が高いことを確認している [12]。北原らは書籍を対象とした推薦システムにおいて、既存の協調フィルタリングにレビューテキストを対象としたトピックモデルを適用することで、書籍の内容を反

映させたアイテム間の類似度を考慮した書籍推薦を実現している [13]。

3. 感情について

感情を抽出する研究において、様々な感情表現を表すために感情軸を感情モデルとして定義している。代表的な感情モデルとして、Ekmanらは人間の感情を「幸福」、「驚き」、「恐れ」、「悲しみ」、「怒り」、「嫌悪」の6種類の基本感情を感情軸として定義している [14]。Plutchikは人間の感情は、「嫌悪」⇔「信頼」、「悲しみ」⇔「喜び」、「驚き」⇔「予測」、「恐れ」⇔「怒り」の8種類の感情軸に分類し、対応する感情が両極となるような対を形成している [15]。

また、中村の感情表現辞典 [16] では、日本語の単語を対象に、感情軸を「喜」、「好」、「安」、「哀」、「厭」、「怖」、「怒」、「恥」、「昂」、「驚」の10軸に分類している。これらに対して山本らは中村の定義する10軸では感情の種類が多く判断しにくいと、ツイートに対して感情分析を行い、結果を基に共起分析およびクラスタリングを行った結果から、Twitterに適した感情軸として「喜・好」、「安」、「昂」、「哀」、「怖」、「怒・厭」、「驚」、「恥」の8軸を定義している [17]。

4. 提案手法

提案手法の概要を図1に示す。提案手法では、山本らの

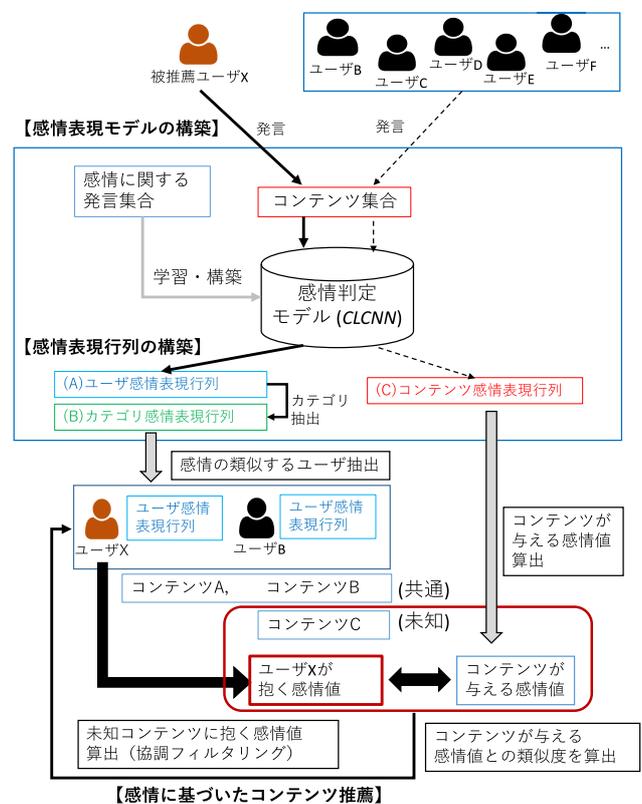


図1 提案手法の概要

Fig. 1 Overview of the proposed method.

8種類の感情軸を用いて、ツイート文からコンテンツに対する感情値を算出するためのニューラルネットワークを感情判定モデルと呼ぶ。ここで、コンテンツに対する感情値は、ユーザがコンテンツに対して抱く感情、およびコンテンツがユーザに与える感情を考慮する。これらの感情を形式的に表すために、下記の3つの感情表現行列を用いる。

(A) ユーザ感情表現行列：対象ユーザがコンテンツに抱く感情を表す行列。

(B) カテゴリ感情表現行列：対象ユーザがコンテンツに抱く感情をカテゴリとして集約した行列。

(C) コンテンツ感情表現行列：コンテンツが不特定多数のユーザに与えている感情を表す行列。

提案方式では、まず、協調フィルタリングを適用することにより、共通のコンテンツに対して類似する感情を持つユーザ群から得られるユーザ感情表現行列に基づいて、あるユーザが未知のコンテンツに対して抱くと予測される感情値を算出し、ベクトル形式でユーザ感情ベクトルとして出力する。また、コンテンツ感情表現行列は、そのコンテンツに対するツイートから算出した感情値を表すコンテンツ感情ベクトルから構成される。あるコンテンツに対して、ユーザ感情ベクトルとコンテンツ感情ベクトルの類似度を算出し、しきい値を超えたものを推薦することで、複数の感情を考慮したコンテンツ推薦を行う。なお、ユーザがコンテンツに抱く感情をカテゴリごとに集約した行列を、カテゴリ感情表現行列として用いることにより、ユーザがツイート文などで言及する共通コンテンツ数が少ない場合において、コンテンツが属する共通カテゴリの類似性に着目したユーザ間距離を算出することもできる。

4.1 感情表現モデルの構築

4.1.1 コンテンツを含むツイートの抽出

ユーザが興味を持つと考えられるコンテンツは、ユーザのすべての発言からより多く発言されている特徴的な単語に現れると考えられる。ユーザが興味を持つと考えられるコンテンツを以下の手順で抽出を行う。

Step-1: マイクロブログから対象となるユーザのツイートを最新 n 件取得する。取得したツイートのうち、リツイートはユーザ自身の感情は考慮されていないため取り除く。

Step-2: コンテンツ名は名詞としてツイート内に現れるため、抽出したユーザの全ツイートに対して形態素解析を行い名詞のみを抽出し、1つの文章としてまとめる。

Step-3: 再構築した文章に対して TF-IDF を適用してコンテンツを表す単語を抽出する。抽出された上位 N 件の単語をユーザが興味を持つコンテンツとして対応付ける。また、各コンテンツは、音楽、ゲームなどの対応するカテゴリと対応付けられる。

4.1.2 感情判定モデルの構築

ユーザがコンテンツに対して抱く感情を抽出するため

表 1 感情の種類と感情語

Table 1 Emotion types and emotion words.

ラベル	感情の種類	感情語
1	喜・好	嬉しい, 面白い, 好む, 憧れ
2	安	ほっと, 安らぐ
3	昂	興奮, 不安定
4	哀	悲しい, 涙ぐむ
5	怖	不気味, 躊躇
6	怒・厭	怒り, 逆上, 嫌い, 悔しい
7	驚	驚く, 放心
8	恥	照れる, 恥じらう

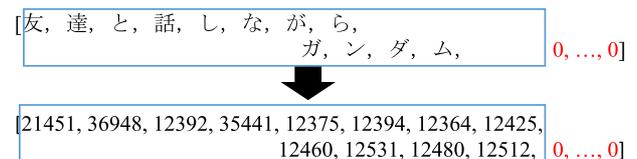


図 2 文字コード (unicode 値) への変換

Fig. 2 Conversion to character code (unicode).

に、ツイートから感情を判定する。本研究では、感情判定モデルに、文章中の単語ではなく、各文字を画像と見なすことで文章を CNN に適用させることができるニューラルネットワークの手法の1つである Character-Level CNN (CLCNN) [18] を用いる。CLCNN を用いる利点として、マイクロブログなどにみられる誤字に強い、学習および判定の際に文章を分かち書きする必要がないという利点がある。

4.1.3 感情判定モデルの学習

感情判定モデルの学習データとして、感情表現辞典に含まれる感情語を含むツイートを抽出し、感情の種類に対応するように1~8の感情ラベルを与えることで学習させる。表 1 にラベルと対応する感情の種類と感情語の例を示す。

感情判定モデルに入力するツイート文の前処理として、ツイートの各文字は文字コード (unicode 値) で符号化される。ツイート文の文字数は最大 u 文字とし、文字数が u 文字に満たない場合は 0 padding で補って調整する (図 2)。提案手法で対象とするツイート文は、最大文字数が $u = 200$ 文字程度の短文を想定している。

感情判定モデルは、符号化された入力ツイート文を4つの convolutional 層で畳み込みを行う。各 convolutional 層では、2, 3, 4, 5 の4つのカーネルサイズのフィルタを適用して畳み込みをする (図 3)。その後、MaxPooling 層で平坦にならした後、全結合層に入力し8次元の感情値に分類する。

4.1.4 コンテンツの感情値の算出

4.1.3 項で学習した感情判定モデルに、4.1.1 項で抽出した日本語ツイート文を入力することで、8種類の感情につ

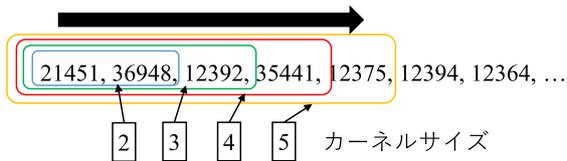


図 3 複数の文字サイズでの畳み込み

Fig. 3 Convolution with multiple character sizes.

表 2 ユーザの感情表現行列の例

Table 2 Example of user emotion expression matrix.

C	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
1	0.38	0.02	0.21	0.05	0.03	0.11	0.17	0.05
2	0.43	0.04	0.21	0.04	0.04	0.16	0.02	0.05
3	0.21	0.07	0.33	0.13	0.04	0.11	0.06	0.03
4	0.26	0.10	0.23	0.08	0.07	0.20	0.04	0.05
5	0.32	0.06	0.26	0.12	0.04	0.11	0.05	0.04

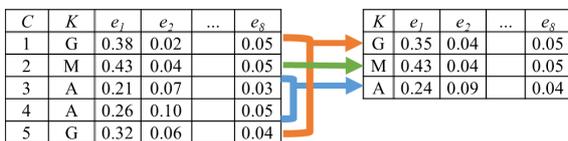


図 4 カテゴリ感情表現行列の構築

Fig. 4 Construction of category emotion expression matrix.

いて感情値 $e_1 \sim e_8$ の算出を行う。感情値 $e_1 \sim e_8$ の値は、活性化関数に softmax 関数を用いて出力結果が合計で 1.0 になるように正規化される。

4.2 感情表現行列の構築

本節では、算出した感情値を用いて以下の 3 つの感情表現行列を構築する。

[(A) ユーザ感情表現行列]

4.1.4 項で判定したコンテンツ名を含むツイートに含まれる感情値 $e_1 \sim e_8$ ごとに平均を算出することで、ユーザのコンテンツ C_x に対して抱く感情値を抽出する。抽出したコンテンツ C_x と 8 種類の感情値 $e_1 \sim e_8$ をユーザの感情表現行列として構築する (表 2)。

[(B) カテゴリ感情表現行列]

さらに、コンテンツ C_x を共通のカテゴリ K ごとにグループ化し、グループ内で各コンテンツ C_x ごとのベクトルの平均値を算出することで、カテゴリ感情行列を作成する。ユーザ感情表現行列からカテゴリ感情表現行列の作成方法を図 4 で示す。図 4 中の表にある K はカテゴリを表しており、 G はゲーム、 M は漫画、 A はアニメといったようにそれぞれカテゴリのラベルを表している。

[(C) コンテンツ感情表現行列]

また、コンテンツがユーザに与える感情を考慮したコンテンツ推薦を行うために、コンテンツが与える感情抽出を行う。そのために任意に抽出した不特定多数のユーザのツ

イートをコンテンツごとに n 件取得し、同様の処理を行うことでコンテンツが与える平均感情表現行列を構築する。

4.3 感情に基づいたコンテンツ推薦

コンテンツ推薦処理では、(1) ユーザが未知のコンテンツに抱く感情値を算出し、さらに (2) コンテンツが与える感情値との類似度を計算することにより実現する。各処理の詳細を下記に述べる。

4.3.1 ユーザが未知のコンテンツに抱く感情値の算出

ユーザの未知のコンテンツに抱く感情値を抽出するには、共通するコンテンツに類似した感情を抱くユーザの未知のコンテンツから抽出する必要がある。ユーザが未知のコンテンツに抱く感情値を、協調フィルタリングを適用し、以下の手順で抽出する。

Step-1: ユーザ A, B の感情の類似度を算出する。ユーザ A, B のユーザ感情行列をそれぞれ $C_a = a_{ie}, C_b = b_{ie}$ ($i = 1, \dots, m, e = 1, \dots, 8$) とする場合、ユーザ A, B 間のコンテンツに頂く感情に関する類似度 $Sim(A, B)$ の値を以下の式 (1) で求める。ここで m はコンテンツ数である。また、各ユーザ感情行列は、各ユーザのコンテンツ数で正規化されている。

$$Sim(A, B) = \frac{\sum_{i=1}^m \sum_{e=1}^8 a_{ie} \cdot b_{ie}}{\sqrt{\sum_{i=1}^m \sum_{e=1}^8 a_{ie}^2} \times \sqrt{\sum_{i=1}^m \sum_{e=1}^8 b_{ie}^2}} \quad (1)$$

Step-2: 感情が似ているユーザが抱く感情値を利用することにより、未知のコンテンツに抱く感情値を算出する。まず類似度がしきい値 θ_{user} 以上のユーザを n 人抽出する。さらに協調フィルタリングを適用することにより、抽出した n 人のユーザのコンテンツへの平均感情ベクトルから、対象ユーザが未知コンテンツに抱く感情ベクトル $UC = [u_{i1}, \dots, u_{i8}]$ を以下の式 (2), (3) により算出する。式 (2) において $Sim(A, B)$ は Step-1 で算出した類似度である。

$$UCx = Sim(A, B) \times [b_{i1}, b_{i2}, \dots, b_{i8}] \quad (x = 1, \dots, n) \quad (2)$$

$$UC = [u_{i1}, \dots, u_{i8}] = \frac{UC_1 + \dots + UC_n}{n} \quad (3)$$

4.3.2 コンテンツ推薦処理

コンテンツ推薦処理では、ユーザが未知コンテンツに抱く感情に対して、コンテンツがユーザに対して与えると期待される感情との類似度を求める。未知コンテンツへの感情ベクトルを $UC = [u_{i1}, \dots, u_{i8}]$ 、コンテンツ感情ベクトルを $V = [v_1, \dots, v_8]$ とし、2 つの感情の類似度 $Sim(UC, V)$ をピアソンの相関係数を用いて算出する。ここで、 \bar{u} は x 個の要素 u_i ($i = 1, \dots, x$) の平均値を、 \bar{v} は x 個の要素 v_i ($i = 1, \dots, x$) の平均値を表す。

$$Sim(UC, V) = \frac{(u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^8 (u_i - \bar{u})^2} \sqrt{\sum_{i=1}^8 (v_i - \bar{v})^2}} \quad (4)$$

$Sim(UC, V)$ がしきい値 $\theta_{contents}$ 以上であるコンテンツ c を、ユーザが未知コンテンツに抱く感情と未知のコンテンツが与える感情が類似するコンテンツであると判断し、ユーザに推薦する。

4.4 意外性の評価

コンテンツ推薦において、似たようなコンテンツばかりが推薦される場合、推薦結果に対するユーザの飽きが生じ、推薦の効果が低下してしまう。このため、ユーザにとって新鮮であり新しい発見のあるコンテンツ推薦を実現するために、意外性について考慮する必要がある。吉川らは、より個人に向けた情報を推薦することで意外性を高めることができるという仮定のもと、Personalizability という個人性を表す指標を設定している [19]。また、村上らは、2種類の推薦モデルを比較し、2つのモデルの推薦結果の差を意外性の高い推薦結果として定義している [20]。

本研究では、推薦されたコンテンツに対して、ユーザ自身が意外であると感じたかを評価することを想定しており、従来の評価手法ではなく、式 (5) のように、推薦されたコンテンツに対してユーザが興味を示したコンテンツ集合のうち、意外だと感じたコンテンツ集合の割合で算出する。

$$\text{意外性} = \frac{\text{意外だと感じたコンテンツの総数}}{\text{興味があるコンテンツの総数}} \quad (5)$$

5. 実験

5.1 感情値判定の評価

5.1.1 目的

提案手法により構築した感情判定モデルを用いて、ユーザのツイートから適切に感情を抽出できるかを確認する。

5.1.2 実験方法

Twitter からランダムに取得したツイート 100 件に対して、想定する種類の感情のラベルを正解データとして付与する。なお、今回対象にした日本語でのツイート文は最大 140 文字であるため、満たないものは 140 文字になるよう 0 でパディングした。提案手法により感情値を算出し、高い感情値を持つ感情を抽出し、正解データと比較することにより、提案手法の感情抽出精度を評価する。

5.1.3 実験環境

表 3 に検証に用いた感情判定モデルの学習に使用したパラメータの詳細を示す。また、実験に用いたツイート文中に含まれる感情表現語および正解感情ラベルの例を表 4 に示す。なお、表 4 のツイート文の例として、 T_1 と T_5 のツイート原文は、それぞれ「キャプ翼の燃えてヒーローのチャンバの意味初めて知った衝撃、、、！」[21]、および「計

表 3 感情判定モデルのパラメータ

Table 3 Parameters of emotion decision model.

パラメータ名	数値
バッチサイズ	1200
使用したモデルのエポック数	2
データ数	1,163,584
精度検証に使うデータ数	100
Twitter からのデータの取得期間	2018 年 6 月~7 月

表 4 ツイート文中の感情表現語と正解感情の例

Table 4 Examples of emotional expressions and correct emotions in tweet.

id	正解感情	感情表現語
T_1	e_7 (驚)	衝撃
T_2	e_5 (怖)	ハラハラ
T_3	e_3 (昂)	ハラハラ, 白熱, 興奮
T_4	e_6 (怒・厭)	このっ, 勝とう, ワンモア, ワンモア
T_5	e_1 (喜・好)	根性, 殺法, 魅力
T_6	e_1 (喜・好)	羞恥, 懐かしい
T_7	e_6 (怒・厭)	救いようが無い?
T_8	e_6 (怒・厭)	文句

表 5 正解判定ツイートの感情値の例

Table 5 Example of emotion value of correct answer tweet.

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
T_1	0.23	0.00	0.00	0.00	0.00	0.01	0.76	0.00
T_2	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
T_3	0.03	0.00	0.65	0.00	0.33	0.00	0.00	0.00
T_4	0.47	0.03	0.09	0.05	0.09	0.22	0.04	0.01
T_5	0.14	0.01	0.13	0.03	0.02	0.65	0.01	0.01
T_6	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.98

表 6 不正解判定ツイートの感情値の例

Table 6 Example of emotion value of incorrect answer tweet.

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
T_7	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00
T_8	0.15	0.31	0.21	0.10	0.09	0.07	0.01	0.06

算より根性なタイプなので、緻密な計算に基づくオダチエン殺法より、ここぞというタイミングでの 10% クリ成立からの大逆転みたいな展開に魅力を感じるお年頃」[22] である。

5.1.4 実験結果

正解判定されたツイート (以下, 正解判定ツイート), 不正解判定されたツイート (以下, 不正解判定ツイート) の例をそれぞれ, 表 5, 表 6 に示す。100 件のツイートに対して, 提案手法による感情抽出精度は, 最も高い感情値で抽出した感情を用いて評価した場合に 87% であった。ま

た、上位3件の感情値で抽出した感情を用いて評価した場合は98%であった。

表5の結果より、正解判定ツイートでは、明確に感情が現れるような単語が文章内に存在することが多いことが確認できる。また、正解ツイートの T_2 、 T_3 は文章の前半の内容がほぼ同じであるが、 T_3 は後半に「興奮した」という単語が存在するため、「怖」の感情よりも「昂」の感情の感情値が高く出力されることが確認できた。同様に、表5の結果から、上位3件の感情値で評価した場合の正解ツイートについて、 T_4 や T_5 については、最も強い感情の次に期待した感情が強く判定されていることから、上位1件では抽出できなかった感情が上位3件に拡張することで感情を抽出することができていることを確認できた。これらの結果より、感情が強く表れるような単語が含まれる文章には正しく感情値を算出することができた。しかし、 T_6 のように「羞恥心」はコンテンツ名を指す単語であるが、「羞恥」に敏感に反応してしまうことで、想定していた「懐かしい」による「喜・好」の感情が非常に低い値で判定されるケースが確認された。このように、ユーザの感情とは無関係な感情単語がツイート文に含まれる場合においては、正しい感情判定が困難になると考えられる。

一方、表6より、不正解判定ツイートの傾向として、 T_7 のツイート文のように「貧困」や「問題」といった「怒・厭」の感情を表す単語が含まれていることがあげられる。ツイート文後半の文章にある「救い」といった単語に敏感に反応してしまうことや、「ない」といった感情の転調を示す単語には対応できないことから、誤った算出結果となると考えられる。また、 T_8 に関しては、ツイート文中に英単語が含まれており、本実験で用いたデータは日本語を対象としていることから、正しい結果を得ることができなかったと考えられる。

5.2 コンテンツ推薦の評価

5.2.1 目的

Twitterのツイート文を対象にした実験を行い、提案手法を用いて、ユーザが抱く感情に類似した感情を与える未知コンテンツの推薦が可能であるかを確認することにより、提案手法の実現可能性および有効性を検証する。

5.2.2 実験方法

以下に示す5つの方式において推薦結果を比較する。評価指標には、適合率、再現率、 F 値、意外性を用いる。意外性の算出は4.4節の式(5)を用いて算出する。

[方式1：ベースライン]

コンテンツ内容に基づいた推薦

コンテンツのベクトル（以下、コンテンツ・ベクトル）として、該当するコンテンツに関する文章をWikipediaから名詞のみを抽出し、TF-IDFで特徴付けしたベクトルを作成する。また、被推薦ユーザのベクトル（以下、ユーザ・

ベクトル）として、ツイート頻度の多かった上位5件のコンテンツに対して、上記のように作成したコンテンツ・ベクトルを足し合わせたベクトルを作成する。コンテンツ・ベクトルとユーザ・ベクトルの類似度をコサイン尺度で算出し、類似度のしきい値 $\theta_{contents}$ 以上であったコンテンツを推薦する。

[方式2：提案手法①（感情マッチングあり）]

ユーザ感情表現行列を用いた推薦

4.3節で示した方式において、ユーザ感情表現行列を用いてユーザ間類似度を算出し、コンテンツ推薦に利用する。

[方式3：提案手法②（感情マッチングあり）]

カテゴリ感情表現行列を用いた推薦

4.3節で示した方式において、カテゴリ感情表現行列を用いてユーザ間類似度を算出し、コンテンツ推薦に利用する。

[方式4：提案手法①（感情マッチングなし）]

方式2において、ユーザが抱く感情とコンテンツが与える感情の感情マッチングをしないで推薦を行う。具体的には、未知コンテンツへの感情値のいずれかが、しきい値 $\theta_{contents}$ 以上のものを推薦する。

[方式5：提案手法②（感情マッチングなし）]

方式3において、感情マッチングをしないで推薦を行う。具体的には、未知コンテンツへの感情値のいずれかが、しきい値 $\theta_{contents}$ 以上のものを推薦する。

なお、各方式のしきい値 θ_{user} 、 $\theta_{contents}$ は、各方式の結果において最も特徴が現れる場合のものを用いる。

実験環境

Twitterにおいて、ユーザのツイート文中に話題の中心として現れる特徴的な単語をコンテンツとして抽出する。コンテンツのジャンルとして、主にゲームやアニメを対象とした。表7に対象ユーザ数および抽出したコンテンツ数などの実験データを示す。また、正解データを作成するために、各方式の推薦結果に対して興味があるコンテンツが推薦されたか、また、興味があるコンテンツは意外なコンテンツであったかについて、被推薦ユーザ本人に対してアンケート調査を実施した。

5.2.3 実験結果

図5、図6、図7、図8に方式1~5の適合率、再現率、

表7 実験データ

Table 7 Experimental data.

項目	抽出数
被推薦ユーザ数	5
推薦候補ユーザ数	50
各ユーザあたりの抽出コンテンツ数	5
各ユーザの対象ツイート数	150
推薦候補コンテンツ数	157
コンテンツ感情表現行列生成のための各コンテンツあたりのツイート数	100

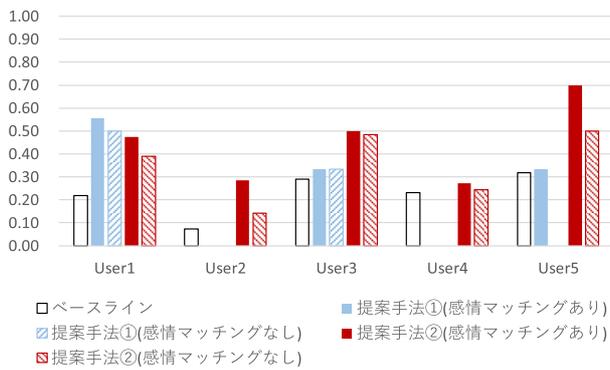


図 5 適合率の評価値

Fig. 5 Evaluation value of precision.

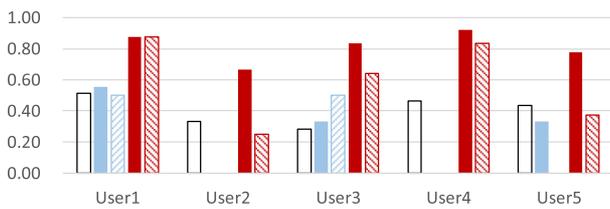


図 6 再現率の評価値

Fig. 6 Evaluation value of recall.

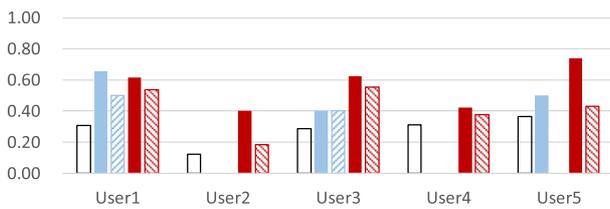


図 7 F 値の評価値

Fig. 7 Evaluation value of F-measure.

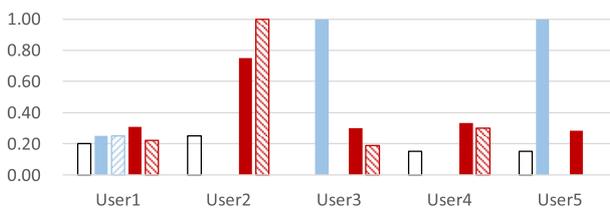


図 8 意外性の評価値

Fig. 8 Evaluation value of serendipity.

F 値および意外性の算出結果を示す。グラフの凡例は図 5 のみに記載した。また、各方式のしきい値として、各方式において最も特徴が現れた下記の値を用いた。

[方式 1] $\theta_{contents} = 0.15$

[方式 2] $\theta_{user} = 0.3, \theta_{contents} = 0.7$

[方式 3] $\theta_{user} = 0.5, \theta_{contents} = 0.7$

[方式 4] $\theta_{user} = 0.3, \theta_{contents} = 0.2$

[方式 5] $\theta_{user} = 0.5, \theta_{contents} = 0.2$

図 5～図 8 の結果より、提案手法②に関する方式 3 と方式 5 において、ベースラインの評価値を上回ることが確認できた。一方、提案手法①に関する方式 2 と方式 4 では、User2, User4 を除いてベースラインの各評価値を上回るこ

とが確認できた。これらの結果より、ベースラインであるコンテンツ内容に基づいて推薦した結果と比較して、ユーザがコンテンツに抱く感情を考慮することで、ユーザの興味を満たすコンテンツ推薦が有効であることが確認できる。

次に、User ごとに感情が一致するコンテンツ推薦ができたかを考察をする。まず User1 は図 5, 図 7 の結果より、ベースラインと比較して、手法①および手法②において他のユーザよりも良い適合率と F 値が得られた。これは、類似ユーザが手法①では 6 名、手法②では 28 名と比較的多くの類似ユーザが得られ、結果として感情が一致するコンテンツが多く推薦されたためである。User1 のように対象ユーザと類似したユーザが多く抽出される場合、感情が一致するコンテンツを推薦できることが分かる。一方、User2～User5 は類似ユーザ数が 0～2 人と少なかったため、手法①では感情が一致するコンテンツを推薦することができなかった。これに対して、手法②では、類似ユーザを抽出するための要素をコンテンツ単体からカテゴリに拡張することにより、より多くの類似ユーザを抽出することができたため、推薦結果を向上することができた。なお、User5 は、手法②においてカテゴリ拡張したにもかかわらず、抽出できた類似ユーザ数が 3 名と少なかった。しかし、User5 の類似ユーザから抽出されるコンテンツ数が 3～5 件と多かったため、User2, User3, User4 の場合よりも多くの感情が一致するコンテンツを推薦することができた。

これらの結果から、ユーザ感情表現行列を用いたコンテンツ推薦では、十分な類似ユーザ数が存在すればより興味を持つコンテンツ推薦が可能であるが、類似ユーザが少ない場合、コンテンツ推薦が行えずベースラインを下回る可能性があることが考えられる。一方、各評価値について、User1 を除いて、提案手法②が提案手法①を上回ることが確認できた。ここで、提案方式①を適用した場合に、User1 は類似ユーザが他のユーザと比較して多く抽出され、この場合において、適合率、F 値が提案方式②の場合を上回ることが確認できた。このことから、十分な類似ユーザ数が存在する場合は、提案方式①を適用した場合の適合率が高くなり、提案方式②を適用した場合では再現率が高くなる傾向があると考えられる。

また、図 5～図 7 の結果において、感情マッチングのあり(方式 2 (提案手法①), 3 (提案手法②))/なし(方式 4 (提案手法①), 5 (提案手法②))の結果を比較すると、提案手法①, ②ともに感情マッチングありのときに、評価値が上昇する傾向にあることが確認できた。この結果は、コンテンツに対してユーザが抱く感情だけでなく、コンテンツが与える感情との関係を考慮することが、ユーザがより興味を持つコンテンツ推薦に有効であることを示している。

図 8 は意外性の結果を示している。いずれのユーザもベースラインと比較して、提案手法の方が意外性のあるコンテンツを多く推薦できていることが確認できた。また、

抽出した類似ユーザについて類似度 θ_{user} がしきい値に近い場合に、その類似ユーザから抽出されるコンテンツが被推薦ユーザにとって意外性のあるコンテンツである傾向があった。これは類似度が低いユーザだけが知っているコンテンツが被推薦ユーザにとって新しい発見のあるコンテンツであることが考えられる。User2 と User4 は、手法②では、類似ユーザの抽出要素をコンテンツ自体からカテゴリに拡張することにより、より多くの類似ユーザを抽出することができたため、多くの意外性のあるコンテンツを推薦できたことが確認できた。一方、User3 と User5 は、手法②では、感情が一致することのできる推薦コンテンツ数も増えたため、結果として意外性を感じたコンテンツの割合が減り、手法①よりも評価値が下がったと考えられる。

6. おわりに

本研究では、ユーザが推薦されたコンテンツに対する感情的な興味を高めることを目的として、コンテンツがユーザに与える感情を考慮したコンテンツ推薦手法を提案し、Twitter を対象とした実験により、提案方式の実現可能性および有効性を検証した。

今後の課題として、提案方式を適用することにより、Twitter や商品レビューサイトと連携して稼働する、実際的な感情を考慮したコンテンツ推薦システムを実現していく予定である。また、感情判定モデルを日本語以外の言語へ対応させることにより、グローバルなコンテンツ推薦の実現も目指していく。

参考文献

- [1] 熊本忠彦, 灘本明代: 共通話題に対する感情的態度の類似度に基づくフォロワー推薦, 電子情報通信学会論文誌 D, Vol.J100-D, No.4, pp.500–509 (2017).
- [2] 山本湧輝, 熊本忠彦, 灘本明代: 話題と感情の可視化に基づくフォロワー推薦, 第 8 回 Web とデータベースに関するフォーラム論文集, Vol.2015, pp.174–181 (2015).
- [3] 加藤慎一郎, 濱皮 礼: Twitter 感情分析を用いた感情値可視化とユーザ推薦システム, 研究報告ヒューマンコンピュータインタラクション (HCI), Vol.2013-HCI-152, No.27, pp.1–8 (2013).
- [4] 村石将嗣, 北山大輔: ツイートの感情推定に基づく対になる感情を誘発する行動の推薦手法, DEIM Forum 2016, P2-2 (2016).
- [5] 佐藤祥子, 高瀬 裕, 中野有紀子: 感情語に基づくことわざ推薦システム, FIT2016 第 15 回情報科学技術フォーラム講演論文集, E-007 (2016).
- [6] 平良浩嗣, 當間愛晃, 赤嶺有平, 山田孝治, 遠藤聡志: 感情推定に基づく小説推薦システムのための認知的評価質問セットを用いたシミュレーション, 情報処理学会第 77 回全国大会講演論文集, pp.585–586 (2015).
- [7] 杉本祐介, 水野忠則, 菱田隆彰: 口コミに含まれる感情語を利用した観光地分類の検討, マルチメディア, 分散協調とモバイルシンポジウム 2014 論文集, Vol.2014, pp.1345–1350 (2014).
- [8] 若宮翔子, 河合由起子, 熊本忠彦, 張 建偉, 白石優旗: 話題に対する多様な感情に基づく Web ページ検索システム, 情報処理学会論文誌, Vol.57, No.1, pp.366–378 (2016).
- [9] 橋本和幸, 中川博之, 田原康之, 大須賀昭彦: センチメント分析とトピック抽出によるマイクロブログからの評判傾向抽出, 電子情報通信学会論文誌 D, Vol.J94-D, No.11, pp.1762–1772 (2011).
- [10] Colneric, N. and Demsar, J.: Emotion Recognition on Twitter: Comparative Study and Training a Unison Model, *IEEE Trans. Affective Computing*, p.1 (2018).
- [11] 服部俊一, 三澤遼理, 石川 博, 高間康史: 価値観モデルを適用した協調フィルタリングによるハイブリッド型推薦手法, 知能と情報 日本知能情報ファジィ学会, Vol.29, No.4, pp.628–636 (2017).
- [12] 西村章宏, 土方嘉徳, 酒田信親: 潜在トピックを利用した協調フィルタリングにおけるトピック情報源の違いに関する調査, 情報処理学会論文誌, Vol.58, No.5, pp.1102–1112 (2017).
- [13] 北原将平, ジェプカ・ラファウ, 荒木健治: レビューテキストを対象としたハイブリッド型推薦手法におけるトピックモデルの有効性について, 第 30 回人工知能学会全国大会, 1K4-OS-06a-4in1 (2016).
- [14] Ekman, P., 工藤 力: 表情分析入門—表情に隠された意味をさぐる, 誠信書房 (1987).
- [15] Plutchik, R.: The nature of emotions, *American Scientist*, Vol.89, pp.344–355 (2001).
- [16] 中村 明: 感情表現辞典, 東京堂出版 (1993).
- [17] 山本湧輝, 熊本忠彦, 灘本明代: ツイートの感情の関係に基づく Twitter 感情軸の決定, DEIM Forum 2015, F5-2 (2015).
- [18] Zhang, X., Zhao, J. and LeCun, Y.: Character-level convolutional networks for text classification, NIPS'15 Proc. 28th International Conference on Neural Information Processing Systems, Vol.1, pp.649–657 (2015).
- [19] 吉川大弘, 森 貴章, 古橋 武: 推薦システムにおける Serendipity (意外性) に関する評価指標, 第 28 回ファジィシステムシンポジウム, pp.286–291 (2012).
- [20] 村上知子, 森紘一郎, 折原良平: 推薦結果の意外性を評価する指標の提案, 2007 年度人工知能学会全国大会 (第 21 回), 2C5-2 (2007).
- [21] 對馬央人: キャブ翼の燃えてヒーローのチャンパの意味初めて知った衝撃,,, 2018 年 6 月 26 日 0 時 54 分 tweet (2018).
- [22] ごった煮@FGO: 計算より根性なタイプなので、緻密な計算に基づくオダチェン殺法より、ここぞというタイミングでの 10% クリ成立からの大逆転みたいな展開に魅力を感じるお年頃, 2018 年 8 月 8 日 11 時 54 分 tweet (2018).

推薦文

本論文では、コンテンツに対する感情をデータベース化し、それらに基づくコンテンツを推薦する実践的な手法を提案している。また、多次元の評価軸に基づくコンテンツを評価する手法も提案しており、膨大なコンテンツの中から、個々の人に対する適切なコンテンツの推薦を目指す意欲的な論文であり、今後の取組において、個々の人の好みを適切に把握するなど、実践的な取組が期待されます。以上の理由により、本論文を推薦いたします。

(マルチメディア通信と分散処理研究会主査 田上 敦士)



藤田 俊

2019年神奈川県立工科大学大学院工学研究科博士前期課程修了。修士（工学）。2019年株式会社スカイウイル入社。情報推薦，パターン認識等のシステム開発に従事。



鷹野 孝典（正会員）

2006年慶應義塾大学大学院政策・メディア研究科博士課程単位取得退学。2006年神奈川県立工科大学助手（現，助教），同大学准教授を経て，2018年神奈川県立工科大学情報学部情報工学科教授。博士（政策・メディア）。人工知能学会会員。情報検索・推薦，マルチメディアデータベースシステム等の研究に従事。