

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} 鷹野孝典^{†2}

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が含まれる発言が多く存在する。複数の感情を考慮することでより共感や興味を抱くコンテンツの推薦が可能であると考えられるが、協調フィルタリングのようにユーザの類似した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おらず不十分であるため、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。提案手法では、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、マイクロブログの一つである 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

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。情報推薦の分野において、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。以上の先行研究を踏まえて、コンテンツ推薦においては、ユーザがコンテンツに対して抱く複数の感情とコンテンツがユーザに対して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。しかし、感情を考慮した情報推薦

は、ポジティブとネガティブの2極値を参照して行うことが多い。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。例えば、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、一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。また、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抱く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を満たすかは判定しない。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。

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、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、コンテンツが与える感情との関係性を考慮した

†1 神奈川工科大学大学院博士前期課程情報工学専攻

†2 神奈川工科大学情報学部情報工学科

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、例えば、あるコンテンツに対して「哀しい」といった感情を抱くと判定されるユーザに対して、そのコンテンツが「哀しい」、「感動」といった感情を与えると考えられる場合などに、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、マイクロブログの一つである 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ー候補として推薦している。加藤ら[3]はユーザ毎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[4]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。佐藤らの手法[5]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。平良ら[6]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、仮定の質問セットと仮定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。

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[7]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、ロコミに含まれる単語とのマッチングを行うことでロコミに含まれる感情を抽出する。若宮ら[8]は Web ページ検索において、検索ワードにより得られた 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ている。橋本ら[9]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。Colneric ら[10]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。

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[11]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。西村ら[12]は協調フィルタリングにトピックモデルを融合させた 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。北原ら[13]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。

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されている。代表的な感情モデルとして、Ekman[14]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。Plutchik[15]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。

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[17]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。

4. 提案手法

本研究では、山本らの8種類の感情軸を用いて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、それぞれユーザ感情行列、コンテンツ感情行列として形式的に表現したものを感情表現モデルと呼ぶ。また、感情表現モデルに基づいて、ツイート文からユーザがコンテンツに対して抱く8種類の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。

提案手法では、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。コンテンツ感情行列の平均を表すベクトルをコンテンツ平均感情ベクトルと呼ぶ。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平均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。提案手法の概要を図1に示す。

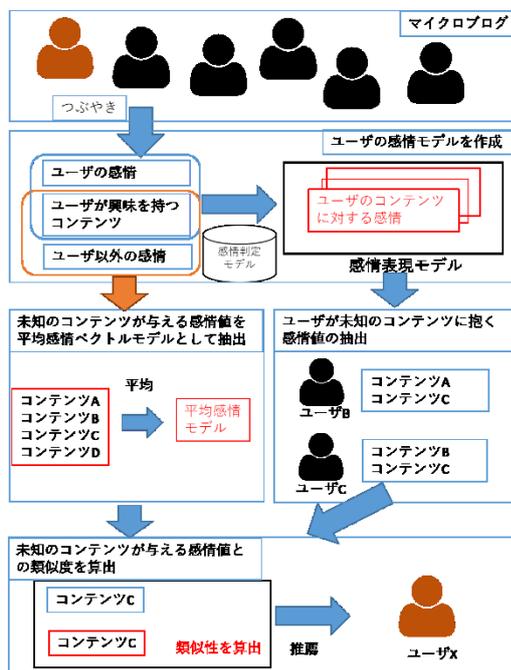


図1 提案手法の概要

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: Twitter API を用いて対象となるユーザのツイートを最新150件取得する。取得したツイートのうち、リツイートにはユーザ自身の感情は考慮されていないため、抽出したツイートから、リツイートは取り除く。

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 Mecab を用いて形態素解析を行い、解析結果から、名詞のみを抽出し一つの文章としてまとめる。Mecab の辞書には、固有の表現に強い Mecab-ipadic-Neologd を用いる。

Step-3: 再構築した文章に対して TF-IDF を用いることで

ンテンツを指す単語を抽出する。算出された上位5単語をユーザが興味を持つコンテンツ C として抽出する。

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には、ツイートから感情を判定させる必要がある。

本研究では、感情判定モデルに、文章中の単語ではなく、各文字を画像とみなすことで、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一つである 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。感情判定モデルの概要を図2に示す。

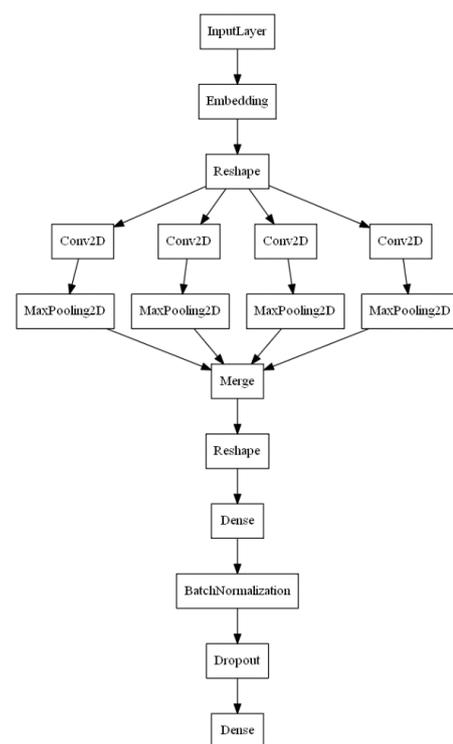


図2 感情判定モデルの概要

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。

表1 感情の種類と感情語

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

学習させる際に、ツイートの各文字を unicode 値に変換することで符号化を行い感情判定モデルに学習させる。本研究で用いる感情判定モデルには、200 文字ごとに 1 文字ずつスライドさせながらツイートを入力させる。ツイートの文字数が 200 に満たない場合は 0 padding を行うことで 200 文字になるように調整をし、200 文字を超える場合には、ツイートの先頭から 200 文字を用いる。本モデルでは 2~5 の複数のカーネルサイズでの浅い畳み込みを行う。これは、2-5-gram を模倣している。畳み込みを行った後、MaxPooling 処理を行い、畳み込み結果を結合し、結合した結果を平坦にならして全結合層にかけ、最終的に 8 つの感情に分類するため、8 次元の値に分類する。

4.1.4 コンテンツの感情の判定

感情判定モデルに 4.1.1 項で抽出したユーザのツイートを与えることで、8 種類の感情値を判定する。判定結果は、活性化関数に softmax 関数を用いるため、出力結果が合計で 1.0 になるように 8 種類の感情 $e_1 \sim e_8$ を抽出する。

4.1.5 感情表現行列の構築

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 $e_1 \sim e_8$ 毎に平均を算出することで、ユーザのコンテンツ C_n に対して抱く感情値を抽出する。抽出したコンテンツ C_n 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 に感情表現行列の例を示す。

表 2 ユーザの感情表現行列の例

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

また、コンテンツがユーザに与える感情を考慮した感情表現モデルを構築するために、推薦候補ユーザ以外のツイートをコンテンツごとに 10 件取得し、同様の処理を行うことで未知のコンテンツが与える平均感情表現行列を構築する。

4.2 ユーザが未知のコンテンツに抱く感情値の抽出

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。ユーザの未知の

コンテンツに抱く感情値の抽出は以下の手順で行う。

Step-1: ユーザの感情表現行列に対して、共通コンテンツの感情の類似度をピアソンの相関係数を用いて求める。例えば、ユーザの A のコンテンツの感情表現モデルを $C_a=[Cae_1, Cae_2, \dots, Cae_8]$, ユーザ B のコンテンツの感情表現モデルを $C_b=[Cbe_1, Cbe_2, \dots, Cbe_8]$ とする場合、コンテンツに対するユーザ A, ユーザ B の同一コンテンツへの感情の類似度を $C_iSim(A, B)$ とする場合、下記の式(1)で求めることができる。

$$C_iSim(a, b) = \frac{(Cae_i - \bar{Cae})(Cbe_i - \bar{Cbe})}{\sqrt{\sum_{i=1}^8 (Cae_i - \bar{Cae})^2} \sqrt{\sum_{i=1}^8 (Cbe_i - \bar{Cbe})^2}} \quad (1)$$

共通コンテンツが複数ある場合は、算出した各コンテンツの類似度の平均を式(2)によって算出することで、ユーザ間の類似度を求める。

$$Sim(a, b) = \frac{C_1Sim(a, b) + \dots + C_nSim(a, b)}{n} \quad (2)$$

Step-2: 類似度がしきい値 θ_{user} 以上の各ユーザのコンテンツの感情値に Step-1 で算出した類似度を乗算することでユーザへの推薦候補となる未知のコンテンツに抱く感情値を抽出する。対象となる未知のコンテンツに複数のユーザの感情値が存在する場合、ユーザの未知のコンテンツに抱く感情値 $UC_i=[UCe_1, UCe_2, \dots, UCe_8]$ は以下の式(3)で算出する。

$$UC_i = \frac{Sim(a, b) \times [Cbe_1, \dots, Cbe_8] + \dots + Sim(a, b) \times [Cne_1, \dots, Cne_8]}{n} \quad (3)$$

算出した未知のコンテンツおよびそれに対して抱く感情値 UC_i をユーザへの推薦候補として抽出する。

4.3 未知のコンテンツが与える感情値との類似度

未知のコンテンツに抱く感情値と未知のコンテンツが与える感情値が類似するコンテンツをユーザに推薦する。

4.2 節で抽出した感情値および 4.1.3 項で構築した未知のコンテンツが与える平均感情表現行列のコンテンツに対する感情値を未知のコンテンツが与える感情値として、類似度を調べるために共通の未知のコンテンツに抱く感情値 $UC_i=[UCe_1, UCe_2, \dots, UCe_8]$ と $AvgC_i=[Avgel, Avgel, \dots, Avgel]$ との類似度をピアソンの相関係数(式(1))を用いて算出する。

算出した類似度がしきい値 $\theta_{contents}$ 以上である上位 5 件のコンテンツをユーザに推薦する。

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、ユーザのツイートから期待した感情を抽出できるかを確認する。複数の感情を考慮した場合、期待した感情が判定されるか確認する。

5.1.1 実験方法

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。提案手法により感情値を算出し、最も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。また、複数の感情を考慮した感情抽出が可能であるかを確認するため、8種の感情のうち上位3件の感情値を持つものに、正解とする感情が含まれているかを判定することにより、複数の感情を考慮した感情抽出精度を評価する。

5.1.2 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。

表3 感情判定モデルのパラメータ

パラメータ名	数値
batch_size	1200
使用したモデルのエポック数	2
データ数	1163584
検証に使うデータ数	100

5.1.3 実験結果

提案手法による感情抽出精度は、最も高い感情値で抽出された感情を用いて評価した場合に87%であった。正解判定されたツイート（以下、正解ツイート）の例を表4に示す。また算出した感情値を表5に示す。また、上位3件で抽出した感情を用いて評価した場合は98%であった。正解ツイートの例を表6に、不正解ツイートの例を表8に示す。また、それぞれについて算出した感情値を表7、9に示す。

表4 正解ツイートの例（上位1件の感情値で評価）

id	正解感情	ツイート
T ₁	e ₇ (驚)	キャブ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、、、!
T ₂	e ₅ (怖)	今日の試合はハラハラする試合だった
T ₃	e ₃ (昂)	ハラハラする試合だった!会場が白熱してて私も興奮した試合だったよ!

表5 正解ツイートの感情値（上位1件の感情値で評価）

	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

表6 正解ツイートの例（上位3件の感情値で評価）

id	想定感情	ツイート
T ₄	e ₆ (怒・厭)	桃「このっ…あああああまた負け

		た!!!!青「このゲームで俺に勝とうなんざ100年早いな!!」桃「ワンモア!!ワンモア!!」
T ₅	e ₁ (喜・好)	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
T ₆	e ₁ (喜・好)	羞恥心流れてる!!!!!!懐かしい!!!!

表7 正解ツイートの感情値（上位3件の感情値で評価）

	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	e ₇	e ₈
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8 不正解ツイートの例（上位3件の感情値で評価）

id	想定感情	ツイート
T ₇	e ₆ (怒・厭)	これだけ取っても借金1000兆円そしてこれだけ取られてもその1000兆円の債権者のほとんどが国民にもかかわらず貧困が問題になるもしかして救いようが無い?w
T ₈	e ₆ (怒・厭)	ウチはブブちゃんの話になってもなー。DEAD ENDのコピーなので。JOEさんが黒いつまみ回しが音が1つ多い方を使ってたけども(セッション事情のボード切替で2台持ってるのか)ブブちゃんはTETSUYA 教則本あやかりのあかべこ…

表9 不正解ツイートの感情値（上位3件の感情値で評価）

	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

表4、5の結果より、正解ツイートでは、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T₂、T₃は文章の前半の内容がほぼ同じであるが、T₃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。表6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、T₄やT₅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

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、また、上位 3 件以内の感情であれば、期待した感情を抽出することが可能であると考えられる。

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、正しい感判定することが困難になると考えられる。

不正解ツイートの傾向としては、 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挙げられる。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、「無い」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つたと考えられる。

5.2 提案手法の実現可能性

5.2.1 目的

Twitter のツイート文を対象にした実験を行い、提案手法によっ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を検証する。

5.2.2 実験方法

下記の 2 つの推薦結果を比較する。

(1)提案手法：4 章で示した方式により、ユーザが未知のコンテンツに抱く感情値と未知のコンテンツが与える感情値との類似性を考慮して、類似度が高かった上位 5 件のコンテンツを推薦する。

(2)ベースライン：ユーザが既知のコンテンツに抱く感情値の平均と未知のコンテンツに抱く感情値との類似性を考慮して、同様に上位 5 件のコンテンツを推薦する。

今回の実験では、各しきい値 θ_{user} , $\theta_{contents}$ をともに 0.7 とする。また、評価指標には、適合率、再現率、F 値を用いる。評価の基準には、推薦結果に対してユーザがコンテンツに抱く感情とコンテンツが与える感情の両方を考慮できているかを確認するため、コンテンツ平均感情ベクトルと未知コンテンツに対するユーザ感情ベクトルの類似度をピアソンの相関係数で算出した平均値を用いた。再現率の指標には、評価指標となるコンテンツの類似度が 0.8 以上であるコンテンツはユーザにとって有益な正解コンテンツとして設定した。

5.2.3 実験環境

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 10 に対象ユーザ数および抽出された対象コンテンツ数などの実験データを示す。表 11 に抽出したコンテンツの例を、表 12 に各ユーザの未知コンテンツを示す。

表 10 実験データ

項目	数
対象ユーザ数	7
抽出された対象コンテンツ数	24
各ユーザのツイート数	150

表 11 各コンテンツの例

コンテンツ id(C)	コンテンツ名
1	FF15
2	FGO
3	コラボ
4	シンフォギア
5	モンスター
6	アリーナ
7	マギレコ
8	ラブライブサンシャイン
9	転生したらスライムだった件
10	デレステ

表 12 各ユーザの未知のコンテンツ(数値:コンテンツ id)

User	C
User1	10, 11, 12, 13, 14, 15, 16, 17
User2	16, 18, 19, 20
User3	1, 3, 4, 5, 14, 15, 16, 17, 21, 22, 23, 24
User4	1, 4, 10, 11, 12, 13, 16, 17, 18, 19, 20
User5	1, 3, 4, 10, 11, 13, 14, 15, 21, 22, 23, 24
User6	3, 5, 6, 7, 8, 9, 14, 15
User7	2, 5, 10, 11, 13, 16, 17

表 13 User4 の既知コンテンツに対する感情行列

C	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
2	0.38	0.02	0.21	0.05	0.03	0.10	0.16	0.05
3	0.21	0.07	0.33	0.13	0.04	0.11	0.07	0.03
5	0.33	0.06	0.26	0.12	0.04	0.11	0.05	0.04
14	0.43	0.04	0.21	0.04	0.04	0.16	0.02	0.05
15	0.26	0.10	0.23	0.08	0.07	0.20	0.04	0.05

表 14 User6 の既知コンテンツに対する感情行列

C	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	e ₇	e ₈
2	0.09	0.02	0.29	0.05	0.04	0.40	0.02	0.09
16	0.15	0.02	0.43	0.09	0.06	0.21	0.01	0.03
18	0.39	0.04	0.18	0.05	0.06	0.20	0.04	0.04
19	0.44	0.03	0.18	0.01	0.01	0.25	0.01	0.07
20	0.26	0.03	0.34	0.03	0.05	0.22	0.01	0.06

表 15 User4 の未知コンテンツに対する感情行列

C	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	e ₇	e ₈
1	0.33	0.02	0.18	0.04	0.03	0.09	0.14	0.04
4	0.23	0.09	0.20	0.07	0.06	0.17	0.03	0.04
10	0.23	0.10	0.16	0.10	0.07	0.07	0.02	0.09
11	0.37	0.04	0.04	0.04	0.15	0.10	0.01	0.08
12	0.38	0.03	0.18	0.04	0.19	0.06	0.02	0.02
13	0.11	0.02	0.20	0.10	0.07	0.24	0.04	0.05
16	0.22	0.03	0.25	0.07	0.07	0.20	0.02	0.04
17	0.44	0.13	0.10	0.04	0.03	0.21	0.01	0.01
18	0.31	0.03	0.15	0.04	0.05	0.16	0.03	0.03
19	0.35	0.03	0.14	0.01	0.01	0.20	0.01	0.05
20	0.21	0.02	0.27	0.02	0.04	0.17	0.01	0.05

表 16 User6 の未知のコンテンツに対する感情行列

C	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	e ₇	e ₈
3	0.27	0.04	0.21	0.08	0.02	0.13	0.02	0.02
5	0.23	0.03	0.21	0.06	0.05	0.16	0.03	0.03
6	0.43	0.07	0.13	0.01	0.13	0.09	0.01	0.01
7	0.07	0.04	0.10	0.07	0.28	0.24	0.01	0.06
8	0.16	0.01	0.09	0.02	0.25	0.33	0.03	0.00
9	0.13	0.01	0.27	0.06	0.34	0.05	0.01	0.01
14	0.17	0.03	0.25	0.05	0.11	0.17	0.01	0.01
15	0.29	0.02	0.32	0.02	0.07	0.06	0.02	0.01

表 17 未知のコンテンツが与える平均感情行列

C	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	e ₇	e ₈
1	0.38	0.03	0.21	0.08	0.10	0.15	0.01	0.02
3	0.31	0.03	0.18	0.09	0.11	0.20	0.01	0.06
4	0.52	0.02	0.12	0.04	0.07	0.09	0.01	0.13
5	0.32	0.16	0.21	0.06	0.07	0.12	0.03	0.03
6	0.40	0.04	0.10	0.02	0.09	0.18	0.10	0.07
7	0.22	0.02	0.16	0.05	0.15	0.23	0.06	0.11
8	0.23	0.02	0.29	0.04	0.15	0.25	0.01	0.01
9	0.35	0.13	0.19	0.13	0.07	0.09	0.01	0.03
10	0.20	0.06	0.22	0.06	0.05	0.21	0.02	0.20
11	0.50	0.12	0.12	0.08	0.08	0.06	0.01	0.04
12	0.34	0.04	0.15	0.09	0.18	0.15	0.04	0.02

13	0.29	0.03	0.16	0.06	0.09	0.15	0.11	0.10
14	0.35	0.07	0.25	0.04	0.03	0.15	0.10	0.02
15	0.28	0.05	0.24	0.07	0.03	0.27	0.02	0.03
17	0.25	0.01	0.15	0.06	0.04	0.38	0.08	0.04
18	0.38	0.07	0.12	0.05	0.06	0.26	0.01	0.04
19	0.51	0.02	0.11	0.08	0.03	0.24	0.01	0.01
20	0.54	0.02	0.14	0.02	0.03	0.24	0.01	0.02
21	0.40	0.05	0.15	0.05	0.05	0.25	0.03	0.02

5.2.4 実験結果

図 3 は、各手法の適合率、再現率、F 値の算出結果をグラフで示している。

図 3 の結果において、提案手法の適合率が共に 1.00 であり、再現率の変化が大きく異なる対象ユーザ User4 と User6 についての結果を中心に考察を行う。対象ユーザ User4 と User6 が既知のコンテンツに対して抱く感情表現行列を表 13, 14, 未知のコンテンツに対して抱く感情表現行列を表 15, 16 に示す。ここで、User4 に感情が類似するユーザ数は 4 名、User6 に類似するユーザ数は 2 名であった。また、未知のコンテンツが与える平均感情行列の一部を表 17 に示す。

表 18, 19 は、User4, User6 について評価指標の類似度とともに、提案手法およびベースラインを用いた場合の推薦結果をランキング化したものを示している。

表 18 User4 への推薦結果

評価指標		提案手法		ベースライン	
C	類似度	C	類似度	C	類似度
18	0.98	18	0.96	4	0.98
19	0.98	19	0.96	16	0.98
12	0.96	12	0.95	20	0.94
11	0.95	11	0.90	18	0.92
1	0.92	1	0.84	19	0.91
20	0.85	16	0.72	13	0.83
16	0.85	20	0.70	17	0.80
4	0.85	4	0.70	1	0.73
17	0.84	17	0.69	10	0.71
10	0.79	10	0.59	12	0.66
13	0.74	13	0.47	11	0.55

表 19 User6 への推薦結果

評価指標		提案手法		ベースライン	
C	類似度	C	類似度	C	類似度
6	0.96	3	0.92	14	0.87
5	0.93	6	0.91	5	0.78
3	0.89	5	0.84	3	0.65

15	0.88	15	0.76	15	0.63
14	0.88	14	0.76	8	0.48
8	0.85	8	0.70	7	0.37
7	0.78	7	0.57	9	0.35
9	0.63	9	0.25	6	0.20

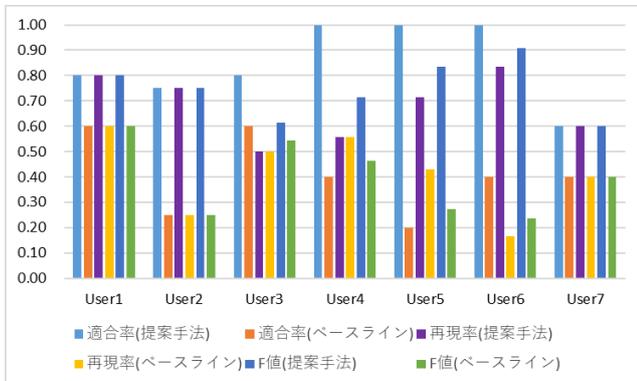


図3 各手法の評価値

User4, User6 に対するコンテンツ推薦において、表 16 より、User4 にはベースラインと比較して提案手法では評価指標においてより類似度が高いと考えられるコンテンツを推薦することを確認できた。User6 においても評価指標の類似度が高いコンテンツをより多くの推薦することが可能であることを確認した。また、どちらのユーザも評価指標の類似度の高い上位 5 件に指定されているコンテンツを全て推薦することができることを確認できた。

また、図 3 の結果より、ほぼすべてのユーザの適合率、再現率、F 値が大幅に上昇していることから、提案手法はベースラインと比較してより高い精度でコンテンツ推薦が可能であることが確認できた。これらの結果から、未知コンテンツが与えるユーザ感情を考慮することによって、ユーザに正しい感情を与えられると考えられるコンテンツを少ない推薦漏れで推薦することができると考えられる。

一方で、User4 の再現率が変化しないことも確認できた。これは、User4 にとって感情がフィットすると、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有益なコンテンツの数が多いため、ベースラインの手法において類似度の高い上位 5 件には指定されていないが評価指標となるコンテンツの類似度が 0.8 以上であるコンテンツを 3 件推薦していることが理由である。このため、有益なコンテンツが多いユーザに対しては、両手法の結果を利用したコンテンツ推薦を行うことでユーザにとって有益であると判断できるより多くのコンテンツの推薦が実現可能になると考えられる。

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、その実現可能性を検証した。

今後の予定として、感情判定手法の精度改善、共通の既知コンテンツを持たないユーザ間の類似度を算出する方法を検討し、コンテンツ推薦手法に適用することで、コンテンツ推薦結果の精度向上を評価していく予定である。

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: ロ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
- [10] N. Colneric, J. Demsar.: Emotion Recognition on Twitter: Comparative Study and Training a Unison Model. IEEE Transactions on Affective Computing, pp.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] P. Ekman, 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] R. Plutchik: 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] X. Zhang, J. Zhao, Y. LeCun: Character-level convolutional networks for text classification. NIPS'15 Proceedings of the 28th International Conference on Neural Information Processing Systems,

