

1. はじめに
2. 既存の手法
3. 感情について
4. 提案手法
5. 数値実験
6. まとめ

[論文紹介] 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

水上 和秀 (Kazuhide Mizukani)
u355020@st.pu-toyama.ac.jp

富山県立大学 工学部 電子情報工学専攻

July 21, 2023

1.1 本研究の背景

2/17

背景

- 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推薦する必要がある
- しかし類似したユーザから推薦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ないため、コンテンツが与える感情との関係を考慮する必要がある

目的

- 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。

2 既存の手法の問題点

3/17

協調フィルタリングを用いたコンテンツ推薦手法

- 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。
- 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。

ポジティブとネガティブの2極値を参照してコンテンツを推薦する手法

- ポジティブとネガティブでは複数の感情を考慮する場合には不十分
- 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。

3 感情について

4/17

本システムで用いる感情の種類

- 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モデルを定義している。
- 感情表現辞典では、日本語の単語を対象に感情軸を「喜」、「好」、「安」、「昂」、「哀」、「怖」、「怒」、「厭」、「驚」、「恥」の 10 軸に分類していた
- Twitter に適した感情軸として本研究では「喜・好」、「安」、「昂」、「哀」、「怖」、「怒・厭」、「驚」、「恥」の 8 種類を定義した

4.1 提案手法の概要

5/17

システムの概要を示す

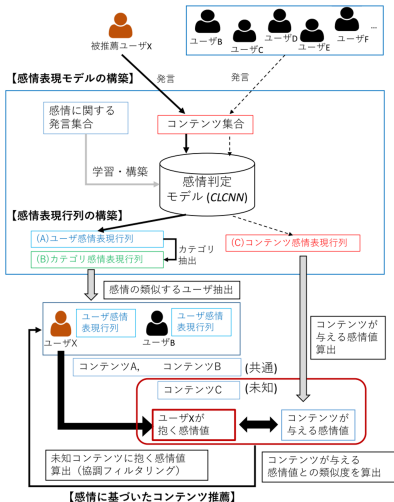


図1 システムの概要

4.1 感情表現モデルの構築 1

6/17

1. 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を以下の手順で抽出する

- 1 Twitter から対象となるユーザのツイートを最新 n 件を取得する
- 2 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
- 3 再構築した文章に対して TF-IDF を適応してコンテンツをあらわす単語を抽出する。

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にツイートから感情を判定する。
- 本研究では各文字を画像とみなすことで文章を畳み込みニューラルネットワーク (CNN) に適応する CLCNN を用いる

4.1 感情表現モデルの構築 2

7/17

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書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

4. コンテンツの感情値の算出

学習した感情判定モデルに、1 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 $e_1 \sim e_8$ の値は合計で 1.0 になるように正規化される。

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

1. はじめに
2. 既存の手法
3. 感情について
4. 提案手法
5. 数値実験
6. まとめ

4.2 感情表現行列の構築

8/17

感情表現行列の構築

本研究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。

1 ユーザ感情表現行列

→ユーザのコンテンツ C_x に対して抱く感情値。抽出したコンテンツ C_x と 8 種類の感情値 $e_1 \sim e_8$ のをユーザの感情表現行表として構築する

2 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

3 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

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

C	K	e_1	e_2	...	e_8
1	G	0.38	0.02		0.05
2	M	0.43	0.04		0.05
3	A	0.21	0.07		0.03
4	A	0.26	0.10		0.05
5	G	0.32	0.06		0.04

K	e_1	e_2	...	e_8
G	0.35	0.04		0.05
M	0.43	0.04		0.05
A	0.24	0.09		0.04

図2 カテゴリ感情表現行列の構築

1. はじめに
2. 既存の手法
3. 感情について
4. 提案手法
5. 数値実験
6. まとめ

4.3 感情に基づいたコンテンツ推薦 1

9/17

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。ユーザの未知のコンテンツに抱く感情値を以下の手順で抽出する。

- 1 ユーザ A、B の感情の類似度を算出するユーザ A、B 間のコンテンツに抱く感情に関する類似度の式を (1) で求める

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

a_{ie} : A がコンテンツ i に対する感情 e の値

b_{ie} : B がコンテンツ i に対する感情 e の値

- 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対象ユーザが未知の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

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

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

4.3 感情に基づいたコンテンツ推薦 2

10/17

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。

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

$V = [v_1, \dots, v_8]$: コンテンツ感情ベクトル

$Sim(UC, V)$ がしきい値 $\theta_{contents}$ 以上であるコンテンツを、ユーザが未知コンテンツに抱く感情と未知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

1. はじめに
2. 既存の手法
3. 感情について
4. 提案手法
5. 数値実験
6. まとめ

4.4 意外性の評価

11/17

意外性の評価

- 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し、意外性を評価する

意外性 = 意外だと感じたコンテンツの総数 / 興味があるコンテンツの総数

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、利用者に対してアンケート調査によって求める

5.1 数値実験 1

12/17

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

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。
- 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

表3 感情判定モデルのパラメータ

パラメータ名	数値
パッチサイズ	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

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

5.2 数値実験の結果

13/17

結果

- ツイート 100 件のうち 87 %は最も高い感情値で評価し、98 %は上位 3 件の感情値で評価できた
- $T_1 \sim T_3$ は最も強い感情が正解の感情となった
- T_4, T_6 は最も強い感情の次に期待した感情が強く判定された。
→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
- T_7 のように「～ない」といった感情の転調を示す単語には対応しにくいことが分かった
- T_8 はツイート文中に英単語が含まれていたため精度が低かった

表5 正解判定ツイートの感情例の例

	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	<u>0.76</u>	0.00
T_2	0.00	0.00	0.00	0.00	<u>1.00</u>	0.00	0.00	0.00
T_3	0.03	0.00	<u>0.65</u>	0.00	0.33	0.00	0.00	0.00
T_4	<u>0.47</u>	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	<u>0.65</u>	0.01	0.01
T_6	0.02	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.98</u>

表6 不正解判定ツイートの感情値の例

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

実験内容

本提案システムの有効性を示すために 5 つの方式において推薦結果を比較する

- 1** コンテンツ内容に基づいた推薦
→コンテンツの内容とユーザのツイート内容の類似度が高いコンテンツを推薦する。
- 2** ユーザ感情表現行列を用いた推薦 (感情マッチングあり)
→ 4.3 で示した方式において、ユーザ感情行列を用いてユーザ間類似度を参照する
- 3** カテゴリ感情表現行列を用いた推薦 (感情マッチングあり)
→ 4.3 で示した方式において、カテゴリ感情行列を用いてユーザ間類似度を参照する
- 4** ユーザ感情表現行列を用いた推薦 (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
- 5** カテゴリ感情表現行列を用いた推薦 (感情マッチングなし)
→方式③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

5.4 実験結果 1

15/17

コンテンツ推薦の評価の結果

適合率 (正解と判定した結果のうち実際の値と一致した割合)、再現率 (実際に正の中からどれだけ正と予測できたかの割合)、F 値 (適合率と再現率の平均)、意外性の算出結果は以下の通りとなった。

各推薦のしきい値

方式①:0.15、方式②:0.7、方式③:0.7、方式④:0.2、方式⑤:0.2

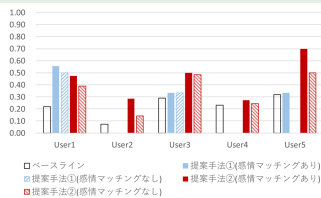


図2 適合率の評価値

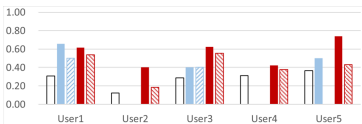


図4 F値の評価値

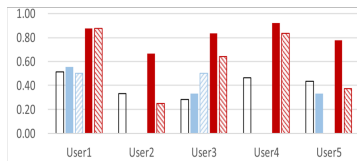


図3 再現率の評価値

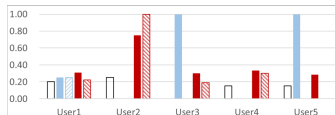


図5 意外性の評価値

1. はじめに
2. 既存の手法
3. 感情について
4. 提案手法
5. 数値実験
6. まとめ

1. はじめに
2. 既存の手法
3. 感情について
4. 提案手法
5. 数値実験
6. まとめ

コンテンツ推薦の評価の結果

- ユーザ感情表現行列を用いた推薦は *User2*、*User4* を除いてベースラインの各評価値を上回ることができた
- カテゴリ感情表現行列を用いた推薦はベースラインの各評価値を上回ることができた。
→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奨が有効であることが確認できた

感情マッチングの比較

- 各推薦において感情マッチングありとなしを比較すると、両手法ともに感情マッチングありの時に評価値が上昇する傾向にあることが確認できた。
- 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。

意外性の評価

- いずれのユーザもベースラインと比較して提案手法のほうが意外性のあるコンテンツを多く推薦することが確認できた。

まとめ

- 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た。
- Twitterを対象とした実験により、提案方式の実現可能性及び有効性を検証した。

課題

- Twitterや商品レビューサイトと連動して、实际的な感情を考慮したコンテンツ推薦システムを実現すること
- 感情判定モデルを日本語以外の言語に対応させること