

ソーシャルメディアおよびECサイトでのレビュー分析のための‘望’感情の抽出 Extraction of "request" intentions for review analysis on social media and e-commerce sites

瀬山 透矢[†]
Yukiya Seyama

Amilcare Astremo[†]
Amilcare Astremo

安達 由洋[†]
Yoshihiro Adachi

1. はじめに

ソーシャルメディアやECサイト上での商品・サービスに関するレビューから改善点や新商品の希望などを記述した文を抽出できると非常に有用である。そして、日本語文から‘要求’や‘要望’を抽出する研究が幾つか報告されている[1, 2, 3]。

一方、著者等は感情語辞書を用いて日本語文の感情分析(分類と検索)を高速に行う感情表現分析システムを開発し発表している。

本研究では、‘望’感情カテゴリと、“が望まれる”や“して欲しい”などのそのカテゴリに属する見出し語を感情語辞書に追加して、希望・要望・要求を記述した文を自動抽出する機能を感情表現分析システムに追加実装した。本発表では、システムの機能概要と実際のレビューデータを用いた性能評価結果について報告する。

2. 感情表現分析システム

我々は授業支援のために自由記述アンケート回答文を高速に分析(分類・検索・可視化など)する技術の研究を行ってきた[4]。授業中に随時アンケートを行い、回答文を高速に分析して、授業に集中できない理由や理解できない理由などの情報を得ることにより、受講生の意見や理解状況を反映した授業を進めることができる。

自由記述アンケート回答文の分析には、話題に基づく分析とともに感情に基づく分析も重要である。そこで我々は、感情語辞書に基づいた高速な感情分析アルゴリズムを開発し感情分析機能を実装した[5]。

我々が開発した感情表現分析システム(Emotion Expression Analysis System, EEAS)は、約7,000語の見出し語をキーとし、10感情カテゴリベクトル[喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚][6]をバリュウとする感情語辞書(Emotion Word Dictionary, EWD)を使用している。EEASは分析対象文から抽出した感情語を用いた説明機構を持っており、この説明に基づいてEWDの語彙やEEASの機能を更新・改良していくことができる。また、4値感情語辞書(FEWD)を用いた4値感情表現分析システム(FEEAS)は、程度副詞も考慮に入れた繊細な感情分析を実行する[7]。

3. ‘望’感情の分類・検索

授業アンケートでは、‘好’や‘厭’の感情だけでなく、“テキストの7ページの式を詳しく説明して欲しい。”や、“おしゃべりがうるさいので注意してください。”などの要求、要望、あるいは希望を記した回答文を高速に分類・抽出できると授業改善に大いに役立つ。また、ソーシャルメディアやECサイトでも、商品・サービスに関するレビュー・クチコミから商品・サービスの改善点や新商品の希

望などを記述した文を分類・抽出できると非常に有用である。例えば、“カメラの解像度がもっと高くなると嬉しいです。”などの文が‘望’感情カテゴリに属する文として分類・抽出できるとユーザの意見として商品開発の参考となる。

3.1 ‘望’感情語

本研究では、EWDとEEASに‘望’感情カテゴリを追加し、‘望’感情カテゴリに属する感情語をEWDに265語登録した。登録した感情語の一部を表1に示す。

表1. ‘望’感情カテゴリの感情語の一部

して欲しい	して頂きたい	して貰いたい
を希望する	を期待する	を望む
が急務です	が先決です	が大事です
あれば有難い	あれば嬉しい	あれば良い
望むことは	期待することは	希望することは

3.2 EWDの作成

感情語辞書管理システム(Emotional-word Dictionary Management System, EDMS)[5]が、‘望’感情を含む[喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚, 望]の各感情カテゴリに属する語データからEWDを自動生成する。

EDMSの感情語辞書生成手順を以下に示す:

- 手順1: 人手で感情ベクトルを付加した感情語 Excel データを入力
 手順2: 感情語を形態素解析器 MeCab で解析して、単語の原形のリストで表現
 手順3: 2 で求めた単語の原形リストを連結して辞書の見出し語を生成
 (単語の原形列を用いることにより、活用形や語形変化に柔軟に対応)
 手順4: 見出し語と感情ベクトルを組みにして EWD Excel ファイルに出力

3.3 EWDの処理例

- 手順1: 感情語データ
 感情語 感情ベクトル
 “して欲しい” [0,0,0,0,0,0,0,0,0,1]
 手順2: 感情語原形リスト
 [‘する’, ‘て’, ‘欲しい’]
 手順3: 原形リストの連結
 “するて欲しい”
 手順4:
 見出し語 感情ベクトル
 “するて欲しい” [0,0,0,0,0,0,0,0,0,1]

[†] 東洋大学総合情報学部

Faculty of Information Sciences and Arts, Toyo University

3.4 EEAS の処理

EEAS は Python の辞書オブジェクト表現 dict を用いて EWD に基づき高速に自由記述文の感情を分析する。

EEAS の感情分析手順を以下に示す：

- 手順 1：EWD を読み込み辞書オブジェクトに格納
(見出し語をキー、感情ベクトルをバリューとする)
- 手順 2：分析対象の自由記述文を形態素解析して単語の原形のリストで表現
- [以下の手順 3 と 4 を、n を 9 から 1 まで減少させながら繰り返す]
- 手順 3：2 で得た文の単語原形リストを先頭から最後まで順に n 語スライスし、切り出した部分リストを連結して EWD の見出し語と照合する。照合した見出し語に対応する感情ベクトルを抽出
- 手順 4：見出し語と照合した感情語の後ろに否定語が続くかリストの照合で調べる。否定語が続く場合は感情カテゴリベクトルを修正
- 手順 5：入力された文から抽出された感情ベクトルのリストを合成 (各対応する成分ごとに最大値を取る)
- 手順 6：5 で得られた感情ベクトルから感情カテゴリ語に変換

3.5 EEAS の処理例

分析対象文：

“テキストの 7 ページの式を詳しく説明して欲しい。”

手順 2：原形のリスト

['テキスト', 'の', '7', 'ページ', 'の', '式', 'を', '詳しい', '説明', 'する', 'て', '欲しい']

手順 3：3 語スライスの連結で見出し語 “するて欲しい” と照合して、感情ベクトル[0,0,0,0,0,0,0,0,1]を抽出

手順 4：後ろに否定語は無い

手順 5：感情ベクトル [0,0,0,0,0,0,0,0,1] を合成

手順 5：感情ベクトルから感情カテゴリ語 ‘望’ に変換

4. ‘望’ 感情処理の性能評価

TSUKUBA コーパス[8]と Google クチコミデータに対して ‘望’ 感情処理の分析精度と計算速度を評価した。

4.1 ‘望’ 感情分析精度

TSUKUBA コーパスに対する分析精度：

混同行列 [[2849 88]
[63 132]]
正解率 0.9517879948914432
適合率 0.6
再現率 0.676923076923077
F 値 0.636144578313253

Google クチコミデータに対する分析精度：

混同行列 [[2969 86]
[19 76]]
正解率 0.9666666666666667
適合率 0.4691358024691358
再現率 0.8
F 値 0.5914396887159533

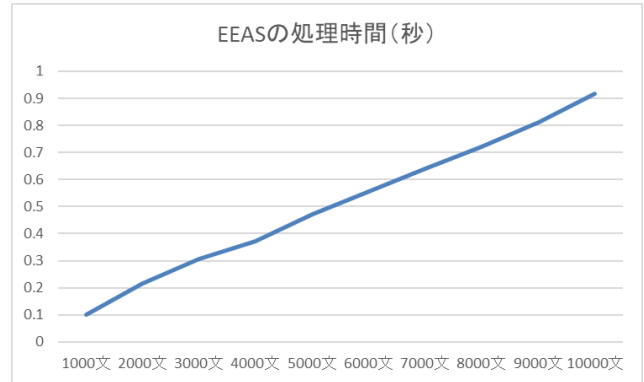


図 1：11 感情カテゴリベクトルの計算時間

4.2 EEAS 計算速度

図 1 に TSUKUBA コーパスに対する 11 感情カテゴリベクトルの計算速度の測定結果を示す。測定環境は Intel Core i7-1068NG7 CPU@2.30GHz, 16 GB memory である。EEAS は 10,000 文の 11 次元感情ベクトルを約 0.92 秒で計算する。この計算時間は 10 感情カテゴリに対する処理時間[5]とほとんど同じであり、EEAS は感情カテゴリや感情語語彙が増えても計算速度にほとんど影響しないことが検証された。

5. おわりに

授業や大規模集会でのアンケート回答文、あるいはソーシャルメディアや EC サイト上での商品・サービスに関するレビューなどから、‘望’ 感情を分類・抽出する研究について報告した。感情語辞書を用いた EEAS による感情分析は高速であり、比較的精度が良い。

今後の課題として、より精度を上げるために感情語データの追加・訓練が必要である。また、近年非常に注目を集めている自然言語処理モデル BERT[9]を用いた日本語文の 11 感情抽出の研究にも取り組み始めている。

謝辞

本研究では、国立情報学研究所の IDR データセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」を利用した。深謝する。

参考文献

- [1] 大塚裕子 他, “自由記述アンケートにおける要求意図判定基準”, 自然言語処理 Vol.11 No.2 pp.21-66 (2004).
- [2] 金山博 他, “要望表現の抽出と整理” (2004).
- [3] 山本瑞樹 他, “文章構造を考慮した自由回答意見からの要望抽出”, (2006).
- [4] Yoshihiro Adachi et al., “Development and evaluation of a real-time analysis method for free-description questionnaire responses”, Proc. the 15th IEEE ICCSE (2020).
- [5] 安達由洋 他, “感情語辞書を用いた日本語文の感情分析”, 可視化情報 Vol.41 No.161 (2021).
- [6] 中村明, “感情表現辞典”, 東京堂出版 (1993).
- [7] 恵谷菜央 他, “感情語辞書と感情分析システム FEEAS”, 情報処理学会第 83 回全国大会(2021).
- [8] 楽天グループ株式会社, “筑波大学文単位評価極性タグ付きコーパス”, 国立情報学研究所情報学研究データリポジトリ (データセット), <https://doi.org/10.32130/idr.2.6> (2014).
- [9] Jacob Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, arXiv:1810.04805v2 [cs.CL] (2018).