

レーダーチャートを用いた感情分析による 商品レビュー評価法

神奈川大学大学院 理学研究科 加納良真 齋藤和巳

1 はじめに

近年、インターネット上のレビューから感情を抽出し分析する研究が注目されている[1, 2]. 特に、商品レビューの文章に含まれる感情を定量的かつ視覚的に表現することは、消費者の意思決定支援やマーケティング戦略の高度化に役立つ. 感情分析分野における体系的な調査として、論文[1]では、感情分析に関する多くのレビュー研究を統合した三次レビュー (tertiary study) により、今後の課題として感情の多次元的表现や可視化の重要性を指摘している. また、論文[2]でも、感情分析の応用や手法を広範に整理し、テキストに含まれる感情の複雑性と可視化の必要性を課題としている.

一方、先行研究[3]では、単語特徴ベクトルに基づく感情ラベル間の類似性や相違性をグラフ構造で可視化する手法を提案しており、感情の構造的な理解と視覚的表現の有効性を示している. また、先行研究[4]では、書き手の性格情報を考慮した感情強度の推定手法を提案し、感情の強度を定量的に扱うことの重要性と有効性を示している. これらの研究はいずれも、従来のポジティブ/ネガティブの二極分類にとどまらない、感情の強度や関係性を捉える多次元かつ視覚的アプローチの重要性を支持している. 本研究では、プルチックの感情の輪に基づく基本感情8種類、混合感情8種類の計16種類の感情を用いてAmazonの商品レビューの文章を分析し、レーダーチャートにより感情とレビュー評点とのAUCスコアを可視化する手法を提案する. これにより、レビューに含まれる多様な感情を視覚的に把握し、商品レビューに対する感情分析可視化法の有効性を示す.

2 提案手法

2.1 感情カテゴリの定義

提案手法では、プルチックの感情の輪に基づく以下の16種類の感情を対象とする.

- 基本感情：

- 喜び (Joy)
- 信頼 (Trust)
- 恐れ (Fear)
- 驚き (Surprise)

- 悲しみ (Sadness)
- 嫌悪 (Disgust)
- 怒り (Anger)
- 期待 (Anticipation)

- 混合感情：

- 愛 (Love) = 喜び + 信頼
- 服従 (Submission) = 信頼 + 恐れ
- 畏敬 (Awe) = 恐れ + 驚き
- 拒絶 (Disapproval) = 驚き + 悲しみ
- 後悔 (Remorse) = 悲しみ + 嫌悪
- 軽蔑 (Contempt) = 嫌悪 + 怒り
- 攻撃 (Aggressiveness) = 怒り + 期待
- 楽観 (Optimism) = 期待 + 喜び

2.2 感情分析可視化法

本研究で提案する可視化手法は以下の手順となる.

1. 前処理として、先行研究[5]の文書データを利用し、先行研究[3]の手法を用いて上記16種類の感情それぞれで各出現単語に対し、各感情での出現度合いをzスコアとして求める.
2. 対象商品のレビュー文書集合を収集し、MeCabを用いて各文書を分かち書きし、16種類の感情それぞれで出現度合いzスコアにより感情スコアを単語平均として求め、感情スコアの大きい順にレビュー文書をソートする.
3. レビュー評価 (星数) が閾値 (実験では4.0に設定) 以上のレビュー文書を正例 (positive)、さもなければ負例 (negative) に設定し、16種類の感情それぞれで感情スコアでソートした文章順に基づきROC曲線を描きAUCスコアを求める.
4. プルチックの感情の輪での16種類の感情配置に基づき、それぞれのAUCスコアでのレーダーチャートを描き可視化する.

本レーダーチャートからは、AUCスコアが基準値の0.5なら、その感情でのレビュー評点は無相関と見なし、0.5より高いか低いかで、正か負の相関を持つとして各商品の感情傾向を分析する.

3 実験による評価

3.1 実験設定

対象はAmazon上の商品レビュー4つの商品（若い女性がよく使う商品2つ，幅広い世代の男女がよく使う商品2つ），各500レビュー（合計2000件）である。

3.2 可視化結果と考察

図1に，各商品に対する感情スコアをレーダーチャートとして示す．なお若い女性がよく買う商品を商品1（青），2（赤），幅広い世代の男女がよく使う商品を商品3（緑），4（紫）とする．実験結果より，4つの商品レビューにおける感情スコアの傾向には明確な差異が見られる．

- 商品1（青）および商品2（赤）は，ポジティブ感情である「Joy, Trust, Anticipation, Optimism」のAUCスコアがいずれも0.5を上回り，特にOptimismは商品1で0.624，商品2で0.666と比較的高い正の相関を示している．一方で，ネガティブ感情では全般的に0.5未満であり，負の相関を示している．すなわち，レビュー文章がポジティブなら評点が高く，ネガティブなら評点が低い一般的な傾向と言える．
- 商品3（緑）は，ポジティブ感情である「Anticipation, Optimism」で0.665, 0.596と高いAUCスコアを示している．一方で，一部のネガティブ感情である「Anger, Contempt」でも，0.479, 0.477と0.5前後またはそれ以上のスコアを示している．すなわち，レビュー文章がネガティブな感情を持ったとしても，レビュー評点とは無相関に近いと言える．
- 商品4（紫）は，ポジティブ感情である「Love:0.689, Anticipation:0.708」などが0.5を比較的大きく超えており，ポジティブ感情との相関が非常に強い商品といえる．これは，ユーザーが本商品に対して一貫して好意的な感情を抱いていることを示している．一方でネガティブ感情である「Anger, Aggressiveness」でも，0.549, 0.620と0.5を上回っており，商品に対するレビューは概ねポジティブであるが，少数ながら強い批判的意見も含まれている．．また，この商品の特徴と見なせる．ただし，Disgustに関しては0.503と，ほぼ0.5に近くレビュー評点とは無相関と言える．

よって，AUCスコア0.5を基準として多次元的に捉えることで，レビューに含まれる感情の種類とAUCスコアを多次元的に捉えることが可能であり，レビュー全体の傾向や特徴を視覚的に把握することができ，レーダーチャートによる感情分析可視化法の有効性が確認できる．

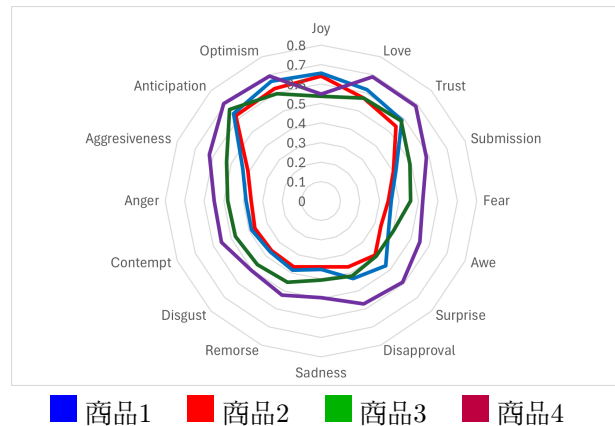


Figure 1: 4つの商品における16種類の感情のレーダーチャート

4 おわりに

本研究では，プルチックの感情の輪に基づき，16種類の感情を用いたレビュー感情分析手法を提案し，レビューに含まれる感情と評点のAUCスコアを多次元的に捉えることが可能であり，レーダーチャートによるレビュー全体の傾向や特徴を可視化した感情分析可視化法の有効性を示した．今後の展望は，BERTなどLLMによる特徴的な単語の埋め込みによるレビュー評価（星数）との相関分析や，より多様な商品カテゴリへの応用，語彙数法への感情分析可視化法の有効性を検証を予定している．

References

- [1] A. Lighthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," **Artificial Intelligence Review**, Vol.54, pp.4997–5053, 2021.
- [2] M. Wankhade, A.C.S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," **Artificial Intelligence Review**, Vol.55, pp.5731–5780, 2022.
- [3] 林咲季, 齊藤和巳, "単語特徴ベクトルとグラフ可視化による感情ラベルの類似性と相違性の分析法", *情報処理学会論文誌*, 2023.
- [4] 鈴木陽也, 秋山和輝, 梶原智之, 二宮崇, 武村紀子, 中島悠太, 長原一, "書き手の性格情報を用いた感情強度推定", *人工知能学会全国大会*, 2022.
- [5] 梶原智之, Chenhui Chu, 武村紀子, 中島悠太, 長原一. 主観感情と客観感情の強度推定のための日本語データセット. *言語処理学会第27回年次大会*, pp.523-527, 2021.