

テキストマイニングを用いた感性評価モデルの構築

Text Mining Approach for Modeling of Kansei Evaluation

(キーワード：印象評価，テキストマイニング，レビューテキスト，LDA)

(Keywords: Impression Evaluation, Text Mining, Review Texts, LDA)

山田篤拓，橋本翔，長田典子（関西学院大学）

1. 背景・目的

感性工学の研究において、特定ドメインの印象軸であるインデックス作成、および各刺激サンプルの印象軸上の得点であるメトリック作成によって、物の感性的な特徴を説明するのに十分な特徴空間である感性評価モデルの構築が行われている。しかし多くの研究法はインデックス作成に被験者実験を伴い、時間や分析にかかる多大なコストが生じている。一方、自然言語処理の分野では、レビューなどのテキストデータからプロダクトに関する情報を抽出する取り組みが行われてきているが、これらの研究で重要視されてきたのは、レビューが「何について（良/悪）」記述されているか、つまり評価対象の識別であった[1]。しかしながら、評価に直接影響を与えるのは、「どう感じたか」という感性的な印象（感性評価モデルにおけるインデックス）であるため、評価対象に対するインデックスを構築することは、特定の印象を与えることを狙ったプロダクトのデザインにおいて有用である。

そこで本研究では、大規模レビューデータから特定のドメインに対するインデックスを作成する手法を提案し、従来のレビューデータ解析手法と比較することで、提案手法と従来手法の差異について論じる。

2. 先行研究

レビューデータからトピック抽出をしている研究としては、金兵らが、商品レビューから自動的に商品ジャンルごとの評価軸を構築する手法を提案している[1]。彼らの手法では、まず「良い」や「満足」など少数の肯定/否定表現を定義し、その周辺情報から評価表現辞書の構築を行う。次に評価表現に係る主節を特徴語として抽出する。そして特徴語間の距離を評価表現と特徴語の組み合わせによって定義し、K-means法を用いたクラスタリングによって評価軸を構築する。彼らの研究における評価軸は、評価に影響する要因となり得るプロダクトの機能や性能、用途に関する軸であり、ユーザがプロダクトに対して感じる印象に関する軸ではない。

3. 提案手法

本研究では前処理、評価語の収集、評価語の選定、評価語空間の推定の4つの手順を踏み、インデックスの作成を行う。

はじめに解析の前処理として形態素解析、表記ゆれの統一を行う。形態素解析にはMeCab[2]を用いて、名詞、動詞、副詞、形容詞、連体詞、接頭辞を用いる。また各単語を原型に変換し、名詞のナイ形容詞語幹は、直前の単語を連結し、接頭詞は次に

続く単語と連結させる。また、商品レビューや大規模なコンテンツなどの審査コメントの自由記述文は、不特定多数のユーザが記述するため、表記ゆれが多く発生する状況にある。表記ゆれはノイズとなり得るため、その解決策として、形態素解析器JUMAN[3]の「代表表記」情報を用いてすべての単語に対して表記ゆれの統一を行う。具体的な手順としては、MeCabにより分割された各単語をJUMANに送り、解析結果の情報に代表表記が含まれる場合に、その単語を代表表記に統一する。

次に評価語の収集は、形態素解析後のいくつかの品詞と、評価に関する複数の辞書を用いて、全ての単語から次の条件に当てはまる単語を評価語として収集する。各条件は、(1)品詞が形容詞自立、(2)品詞が名詞の形容詞語幹、(3)評価表現辞書[4]に含まれる、(4)日本語評価極性辞書（用言編，名詞編）[4,5]中の「評価」の属性が付与されている単語に含まれる、である。

次に評価語の選定について述べる。評価表現には評価対象に対する評価者の感情や感情を表す行為を示す表現である「内評価」と、評価対象の特徴を表す表現である「外評価」という2つのカテゴリが存在する[6]。本研究では評価対象の与える印象のトピック抽出が目的であるため、評価者の状態を示す内評価の単語は不必要だと考えられる。そのため、評価表現に対して内評価/外評価の属性が付与されている日本語評価アプレイザル辞書[7]を用いて、収集した評価語のうち内評価の属性が付与されていない評価語のみを選択する。また極端に頻度の高い/低い単語の存在はノイズとなるため、50%以上のプロダクトのレビューに存在する評価語と0.1%以下のプロダクトのレビューにしか存在しない評価語は取り除く。

最後に評価語空間の推定について述べる。評価語空間の推定には、テキストデータのクラスタリングに広く用いられるLDAの拡張版であるHDP-LDA[8]を用いる。LDAではあらかじめ分析者がトピック数を定める必要があるが、HDP-LDAではトピック数が自動で決定されるという利点がある。事後分布の近似手法としてはCGS法を用いて、サンプリング回数500回とする。ハイパーパラメータは、 $\alpha=2.0$ 、 $\beta=0.3$ 、 $\gamma=0.01$ とする。

4. 腕時計のレビューデータ解析と先行研究との比較

本章では提案手法を金兵らの先行研究で用いられたデータに適用し、その結果を比較し考察する。

使用データは、先行研究と同じ楽天株式会社が提供する公開データセットである「みんなのレビュー・口コミ情報」の腕時計に対するレビューデータのうち、2010年から2013年までの

表 1 腕時計の商品レビューに対する先行研究の評価軸構築結果と本研究のインデックスの作成結果の比較

	ラベル	軸/トピックを構成する重要単語
先行研究	価格	お値段, 価格, 値段, 金額, とこ, 送料, 種類, カラーバリエーション
	文字盤	針, 見た目, ケース, 大きさ, 箱, 文字盤, 画面, 日付, 形, 色合い
	外装	バンド, ベルト, 部分, 液晶, 本体, 最初, 作り, 厚み, 革, 仕上げ
	つけ心地	重さ, つけ心地, 安っぽさ, 装着感, つけ心地, 説明書, 手間, チタン, 使用感, 返信, 重量
	品質	使い勝手, 質, 質感, 評判, フィット感, 使い心地, 品質, 重量感, コストパフォーマンス, 性能, 相性, 視認性, 精度
	音	音, 動作音
本研究	プレゼント	主人, 子供, 人, デザイン性, 娘, モデル, 息子
	重厚な高級感	高級だ, 重い, 重厚だ, 質感, 非常だ, シッカリ
	丁寧で迅速	丁寧だ, 迅速だ, 素敵だ, 最高, おしゃれ, カジュアルだ
	値段相応	安っぽい, 相応, おしゃれ, 堅い, ビックリ, 違う
	軽くて見やすい	軽い, 見易い, 薄い, 簡単だ, 正確だ, 不要だ
	見づらい	ライト, 見難い, 暗い, カッコイイ, 見辛い, 問題ない
	子供用にタフで丈夫	子供, タフだ, 丈夫だ, 迅速だ, 丁寧だ, お買い得
	上品で素敵	上品だ, 素敵だ, 小さい, 華奢, 可愛らしい, 長い
	おしゃれでカラフル	見易い, 子供, 違う, カラフルだ, 汚れる, おしゃれ
	アンティーク	アンティーク, ダーク, ナチュラルだ, 馴染む, 堅い, 違う
	軽くて子供に合う	軽い, 子供, 邪魔, 小さい, 合わせる, 違う
	面白くて変わった	面白い, おしゃれだ, 変わる, 珍しい, 光る, 明るい
	目覚ましに便利	目覚まし, スッキリ, 気付く, 便利だ, 悪い, 浅い
	高度で本格的	高度, 本格, ハイだ, 落ちる, 満載, 締める
—	ワイドだ, 無垢だ, 平均, 潤滑, 最良だ, 真剣	
無骨でタフ	タフだ, 無骨だ, 剛健, 正確だ, 質実だ, 黄色	

3年間分のレビューを使用し、先行研究と同様に「レビュータイトル」と「レビュー内容」を連結させて1つのレビューとして扱った。その結果、レビュー数は203,086件、総単語数4,514,325、総語彙数は19,689語となった。

評価語の収集・選定の結果、評価語として収集された単語は3,097語となり、そのうち内評価の単語は493語、インデックスの作成に用いる内評価でない評価語は2,604語となった。

これらのデータをもとにインデックスを作成した結果、15のトピックが抽出された。それらのトピックと各トピックのterm-score上位6語、および比較対象として先行研究の結果を表1に示す。ラベルは軸/トピックを構成する重要単語から解釈して付与した。表からわかるように、従来手法ではそのドメインにおいて「ユーザが何について評価しているか」を示しているのに対して、提案手法では「どう感じたか」という感性的な印象が抽出されていることがわかる。このような対象の与える印象のパターンは従来法では得られなかった知見である。これらの手法を組み合わせることによって「何についてどう感じたか」がより明確になるため、テキストデータからの感性情報の抽出に有効であると考えられる。

5. まとめ

本研究は大規模レビューデータからインデックスを作成するために、複数の辞書を用いて評価語を収集・選定し、HDP-LDAを用いて評価語空間の推定を行った。その結果、先行研究とは異なったインデックスを作成することができ、それらのトピックの有用性が示唆された。従来手法のような評価対象のトピックと本手法の印象評価のトピックを組み合わせることで、レビューデータからより詳細な感性情報の抽出ができる

ことが期待される。

参考文献

- [1] 金兵裕太, 沼尾雅之.: ネットショッピングサイトの商品レビューを利用したジャンル毎の評価軸の自動構築とその応用, DEIM Forum, 2016.
- [2] Kudo, T., Yamamoto, K., & Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphological Analysis, Proceedings of EMNLP 2004, pp. 230–237, 2004.
- [3] Kurohashi, S.: Improvements of Japanese morphological analyzer JUMAN, Proceedings of the Workshop on Sharable Natural Language Resources, pp. 22–28, 1994.
- [4] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, & 福島俊一: 意見抽出のための評価表現の収集, 自然言語処理, 12, 3, pp.203–222, 2005.
- [5] 東山昌彦: 述語の選択選好性に着目した名詞評価極性の獲得, 言語処理学会 第14回年次大会, 2008.
- [6] 佐野大樹: 『日本語アプレイザル評価表現辞書(態度表現編)』の構築～評価の多様性を捉えるための言語資源の開発～, 言語処理学会 第17回年次大会発表論文集, pp.115–118, 2011.
- [7] 佐野大樹: 日本語アプレイザル評価表現辞書-態度評価編-JAppraisal 辞書 ver1. 言語資源協会, 2011.
- [8] Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M.: Sharing Clusters among Related Groups: Hierarchical Dirichlet Processes, Advances in Neural Information Processing Systems 17, pp. 1385–1392, 2005.