

商品レビューにおける評価極性変化の同定

荒井 利輝[†]高村 大也^{‡§}奥村 学[‡][†] 東京工業大学[‡] 東京工業大学科学技術創成研究院[§] 産総研 AIRC

arai@lr.pi.titech.ac.jp, {takamura,oku}@pi.titech.ac.jp

1 はじめに

インターネットの普及により、一般のユーザが商品に対して手軽にレビューを書けるようになった。それに伴いユーザの意見や評価を分析する研究が盛んに行われてきている。そしてユーザが商品に対して肯定的な評価をしているのか、もしくは否定的な評価をしているのかを判別する評価極性同定の研究も数多くなされている [1]。一方で、レビュー中には

「この小説の広告を見て面白そうだと思って買った。けどつまらなかった。」

といった記述が存在する。この記述ではまず商品 (小説) に対して購入前などの事前評価として「面白そう」と肯定的な評価 (予想) をしている。そして、その評価が「広告」という情報源を根拠にしていることがわかる。また「面白そう」という事前評価に対して、購入後などの事後評価が「つまらなかった」と、事前評価が悪い方向に変化したことがわかる。以上のことはマーケティングに役立ち、また購入の参考にもなる。マーケティングに関しては「広告」によってユーザが購入したことがわかり、宣伝効果があったことが確認できる。また「広告」を見て「面白そう」と考えて購入を検討しているときに、自分と同じように「広告」を見て「面白そう」と考えて購入したユーザの事前評価が悪くなったことがわかれば、購入の参考になる。

よって情報源や事前評価の極性変化を明らかにすることは有益であり、これらを同定したい。そこで本研究ではユーザが事前に評価を行っている文を初めに同定し、その次に事前評価の情報源が何であるか、またその事前評価がどのように変化したか判別する。これらの判別の問題を定義するとともに、その問題を解く手法を提案する。本研究では上の問題をそれぞれ文分類問題として定式化し、有用と思われる素性を考案した。そして実験の結果、それぞれの手順において考案した素性の有用性を確認することができた。

2 関連研究

既存の評価極性同定の研究 [1] は、記述の評価極性を判別することに主に焦点が当てられており、レビュー中に事前評価が記述されているか、事前評価がどのように変化したか、さらに、事前評価の情報源は何かといった点に焦点を当てたものは我々の知る限り存在しない。しかし関連する研究はいくつか存在する。

1 節の例では、「広告」という情報源をきっかけに「面白そう」と考えて購入したことがわかる。つまり本研究が扱う問題は、購入理由に関して分析を行っている側面があるといえる。商品の購入理由を分析した研究として、石野ら [2] が挙げられる。石野らは商品レビューを対象に「たくて購入」などの語彙構文パターンを活用して購買意図の抽出と分析を行った。また購入理由以外に、ユーザの購入前の行動を取り扱った研究が存在する。Adelら [3] は Twitter のユーザの購買行動に対し注意、関心、欲求、行動 (購入) などの段階を定義し、段階の特定を行った。しかし、これらの研究は事前評価の情報源に関する分析を行っておらず、本研究とは異なる。

また 1 節の例では「面白そう」から「つまらなかった」と極性が悪い方向に変化していると述べたが、これは極性が反転したともいえる。極性の反転に関して分析した研究として、井上ら [4] が挙げられる。井上らは質問サイトから赤ちゃんポストに関する意見の収集を行った。特に「赤ちゃんの命が救われるので賛成です。ただ、ネーミングは悪いと思います。」の下線部を極性反転文と呼び、極性反転文を検出する手法を提案した。検出する際には極性反転の表現を分析し、規則を作成することで検出を行った。井上らが挙げた例では評価極性は反転しているが、ある対象に対する評価極性が変化したわけではない。そのため、本研究と分析対象が異なるといえる。

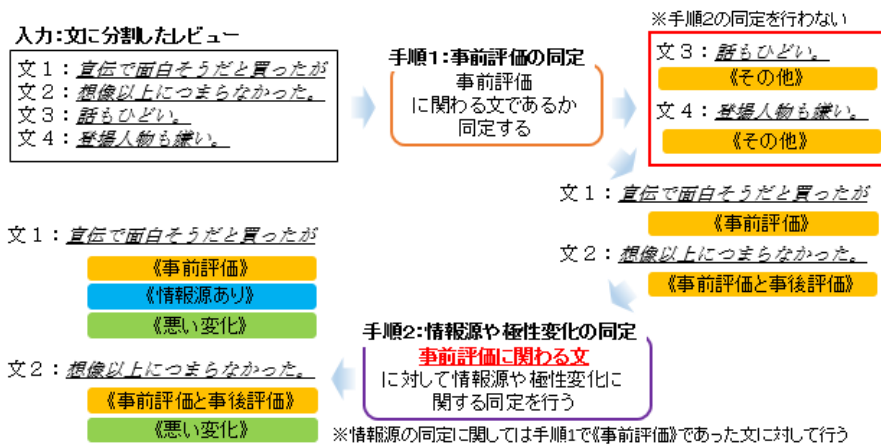


図 1: 手法の概要

3 ラベルの定義

レビューにおける評価極性変化を捉えるために、予め分割した各単文¹(以下、文と呼ぶ)に対し、次のようなラベルを考える。まず、各文は《事前評価》,《事前評価と事後評価》,《その他》のいずれかの分類クラスに属するものとする。図1の文1にある「面白そう」のように、その文中に事前評価のみが記述されていれば《事前評価》, 図1の文2にある「想像以上につまらなかった」のように、事前と事後評価の両方が記述されていれば《事前評価と事後評価》となる。

事前評価が記述されている《事前評価》に属する文については、その事前評価の情報源に関する記述について《情報源あり》および《情報源なし》というクラスを考える²。販売や生産に関わる企業や、知人や商品のユーザなどの第三者が発信した情報源に関する記述があれば《情報源あり》となる。例えば図1の文1に「宣伝」とあるので《情報源あり》とわかる。

同様に、事前評価が記述されている《事前評価》と《事前評価と事後評価》の2クラスに属する文については、事前評価と事後評価の評価極性の違いについて、《良い変化》,《変化なし》,《悪い変化》のいずれかのクラスを考える。事前評価よりも事後評価が良い方向に変化していれば《良い変化》, 逆に悪い方向に変化していれば《悪い変化》, それ以外であれば《変化なし》とラベル付けする。例えば図1の文1の「面白そう」と文2~4の評価を比較すると《悪い変化》であることがわかる。文2も「想像」よりも「つまらなかった」から《悪い変化》とわかる。

¹「。」などの記号や「けど」,「が」などの接続助詞の他、様々な手がかり表現に着目してレビューを単文に分割した。

²本来であれば《事前評価と事後評価》にも情報源の記述の有無が考えられるが、今回使ったデータではそのような文が非常に少なかったので無視することにする。

4 提案手法

本研究では予め分割した文に対して、事前評価に関わっている文かどうかを同定し、その次に情報源や評価極性の変化に関する同定を行う。手法の概要を図1に示す。なお以後、評価極性を極性、評価極性の変化を極性変化と呼ぶ。

手順1では全ての文に対して《事前評価》,《事前評価と事後評価》,《その他》のどれであるかを判定する。手順2の情報源の同定では、手順1で《事前評価》であった文に対して《情報源あり》,《情報源なし》どちらであるかを判定する。手順2の極性変化の同定では、手順1で《事前評価》か《事前評価と事後評価》であった文に対して《良い変化》,《変化なし》,《悪い変化》のどれであるかを判定する。なお分類器にはSupport Vector Machine (SVM) を用いる。

4.1 提案素性

本節では各分類における素性を表1をもとに説明する。なお基本素性には単語の表層形, 原形, 品詞タグの uni-gram, bi-gram, tri-gram の頻度を用いる。

(1) 手順1(素性1-4)

素性1と2は、文と単語の位置情報に関する素性である。事前評価(購入前)から事後評価(購入後)という流れの記述が多いため、事前評価に関わる文や単語(「広告」など)³はレビューの前半に出現しやすい。素性3と4は評価語を用いている素性である。ある評価を抱いて購入したり、ある評価をきっかけに商品を手にとったりするケースが多い。例えば「楽しみにして購入。」や「面白そうなので読んだ。」などがある。

(2) 《事前評価》の情報源(素性5-6)

情報源があるときは情報源に関する単語(「広告」

³実験で使用したコーパスとは別のものを各ドメインで用意し、Yihら[5]の定義した計算式を用いて名詞の出現位置の計算を行っている。その結果を素性値として使用している。

素性	
1	最初と最後から数えた時、レビュー全体を5分割した時の文の位置
2	分類対象の文にある名詞の出現位置の最小値, 平均, 最大値
3	評価語⇒「購入」or「買う」(左の順であるか)
4	評価語⇒「から」or「ので」or「ため」⇒過去の時制(左の順であるか)
5	“情報”, “批評”カテゴリの単語があるか
6	“情報”, “批評”カテゴリの単語⇒評価語(左の順であるか)
7	分類対象の文の極性値が+1 or -1, かつ手がかり表現があるか
8	手がかり表現がある文と次の文の極性値を比べて大きい or 同じ or 小さい
9	総合的な評価の極性値が+1 or -1 であるか
10	分類対象の文と総合的な評価の極性値を比べて大きい or 同じ or 小さい
11	ポジティブ, ネガティブな評価語が両方あるか
12	文末に一番近い評価語の極性が+1 or -1 であるか
13	素性9と同じ

表 1: 素性一覧

手がかり表現	
が, けど, けれど, けれども, のに, だけに, ものの	

表 2: 極性の反転を表す手がかり表現

など) が出現している傾向があり, 素性5と6はその点に基づいている素性である. 日本語語彙大系 [6] のカテゴリから“情報”(単語数: 277)と“批評”(単語数: 51)を選択し, それらに属する単語を情報源に関する単語として用いている. 例えば「広告を見て買った。」や「口コミで期待して読んだ。」などが素性5と6に対応する.

(3) 《事前評価》の極性変化(素性7-10)

素性7と8は「けど」など, 極性の反転を表す手がかり表現に着目した素性である. 「面白そう」など事前評価に関する記述がある文に, 極性の反転を表す手がかり表現があるときは極性変化に関わることが多い. 表2は本研究で用いている手がかり表現である.

例えば図1の文1は「面白そう」と「けど」より事前評価が悪い方向に変化した可能性が高いと考えることができる. 素性7はこちらに対応する. また事前評価に関する記述であり, なおかつ手がかり表現がある文の次にある文は事後評価に関わる記述であることが多い. そのため2つの文の極性を比較することで極性変化がわかることが多い. 例えば図1の文1と, その次の文2の極性を比較すると《悪い変化》だとわかる. こちらは素性8に対応する.

素性9と10は分類対象の文以外の文の極性値の合計をユーザの総合的な評価(事後評価)として活用している素性である. 素性9は総合的な評価と極性変化の極性は一致することが多いという点に基づいている素性である. 例えば図1では文2~4の記述から総合的な評価はネガティブとわかり, 《悪い変化》と極性が一致している. また事前評価に関わる文1と総合的な評価の極性を比較することで《悪い変化》だとわかる. こちらは素性10に対応する.

(4) 《事前評価と事後評価》の極性変化(素性11-13)

	事前	事前と事後	その他
小説	1530	357	60138
掃除機	1184	636	34802
イヤホン	1137	484	39696

表 3: 手順1のデータ

	事前	
	情報源あり	情報源なし
小説	436	1094
掃除機	334	850
イヤホン	390	747

表 4: 手順2(情報源)のデータ

	事前			事前と事後		
	良い	なし	悪い	良い	なし	悪い
小説	370	606	554	123	51	183
掃除機	314	676	194	302	134	200
イヤホン	242	575	320	211	115	158

表 5: 手順2(極性変化)のデータ

素性11は「期待した内容ではなくいまいち」といった事例に対応している. 《事前評価と事後評価》で極性が変化する場合はポジティブ, ネガティブな単語が両方あることも多く, この例でも「期待」, 「いまいち」の両方がある. 素性12は文末の評価語に着目した素性である. 総合的な評価だけではなく, 《事前評価と事後評価》の文中にある事後評価と極性変化の極性も一致しやすい傾向がある. そして文末に事後評価があることが多く, 図1の文2でも文末の評価語は「つまらなかった」であり, 《悪い変化》と極性が一致する. 素性13は素性9と同じ理由で組み込んだ素性である.

以上全ての素性作成の際には MeCab⁴と CaboCha⁵を使用している. 極性値の計算には日本語評価極性辞書 [7][8]を使用し, 文の極性値は, 文にある評価語の極性値の合計が+1以上で+1, -1以下で-1としている. また総合的な評価の極性値は, 分類対象の文以外の文の極性値の合計が+1以上で+1, -1以下で-1としている. そして評価語と「ない」, 「まい」, 「ん」, 「ず」, 「ぬ」といった否定表現(全て原形)が同じ文節にある, または評価語がある文節の係先に否定表現がある場合は極性値を反転させる.

5 実験

5.1 データセット

Amazon⁶における3つのドメイン(「小説」が5,691件, 「掃除機」が4,764件, 「イヤホン」が4,771件)の

⁴<http://taku910.github.io/mecab/>

⁵<http://taku910.github.io/cabocha/>

⁶<https://www.amazon.co.jp/>

		Base	All
小説	事前	0.630	0.663[†]
	事前と事後	0.636	0.653
掃除機	事前	0.670	0.681
	事前と事後	0.820	0.821
イヤホン	事前	0.699	0.725[†]
	事前と事後	0.736	0.747

表 6: 手順 1 の実験結果

	Base	All
小説	0.668	0.683
掃除機	0.834	0.869[†]
イヤホン	0.864	0.873

表 7: 手順 2(《事前評価》の情報源) の実験結果

レビューを文に分割し, 作業員 1 名によってアノテーションを行った. 表 3-5 がその詳細である.

5.2 実験設定

4.1 節の基本素性のみを用いたものをベースラインとする. SVM(線形カーネル) のツールとしては scikit-learn⁷ を用いており, 多値分類器を構築する際には one-versus-rest 法を採用した. 実験結果の評価には各クラスに対する F 値を用いた. ハイパーパラメータ C は 0.1 から 1 の間で 0.1 刻み, 1 から 10 の間で 1 刻みに変化させ, 訓練データにおいて 5 分割交差検定を行った結果, F 値の高いものを選択した. また手順 2 の実験では手順 1 のラベルが全て正しく特定されたものとして, 5 分割交差検定を行った.

5.3 実験結果

ベースライン (Base) と, それぞれの手順の提案素性を全て追加した場合 (All) の結果を比較した. 表 6-9 の数値は各クラスの F 値の値である. † はベースラインの結果と提案素性を全て追加したときの結果の差が, 並び替え検定において統計的に有意⁸であることを示す. 実験結果より, ほとんどのクラスにおいて提案素性が有効であることを確認できた.

6 まとめ

本研究では情報源と評価極性の変化の同定に取り組み, 有用な手法を提案した. 今までの研究において, 本研究のような取り組みに焦点を当てたものはなく, 本研究は新規性の高いものであると考える. そして実験の結果, 考案した素性の有用性を確認することができた. しかし本研究で考案した素性以外に, それぞれの手順において有用な素性があると考えている. 今後

⁷<http://scikit-learn.org/>

⁸有意水準は 0.05 とした.

		Base	All
小説	良い変化	0.621	0.655
	変化なし	0.570	0.597[†]
	悪い変化	0.559	0.616[†]
掃除機	良い変化	0.739	0.745
	変化なし	0.824	0.808
	悪い変化	0.412	0.491
イヤホン	良い変化	0.664	0.710[†]
	変化なし	0.719	0.739[†]
	悪い変化	0.484	0.543[†]

表 8: 手順 2(《事前評価》の極性変化) の実験結果

		Base	All
小説	良い変化	0.840	0.837
	変化なし	0.582	0.619
	悪い変化	0.876	0.884
掃除機	良い変化	0.831	0.846
	変化なし	0.884	0.888
	悪い変化	0.739	0.774[†]
イヤホン	良い変化	0.858	0.873
	変化なし	0.887	0.889
	悪い変化	0.792	0.811

表 9: 手順 2(《事前評価と事後評価》の極性変化) の実験結果

はエラー分析をもとに新たに有用な素性を考案していきたい. また今回使用したドメイン以外に新たなドメインを追加し, 本研究で考案した素性の有用性の確認をしていく.

参考文献

- [1] Bing Liu. “Sentiment Analysis and Opinion Mining”. Morgan & Clapool, 2012.
- [2] 石野亜耶, 村上浩司, 関根聡. “商品レビューからの購買意図の抽出とそれを用いた商品検索システムの構築”. 言語処理学会第 20 回年次大会発表論文集, pp.622-625, 2014.
- [3] Heike Adel, Francine Chen and Yan-Ying Chen. “Ranking Convolutional Recurrent Neural Networks for Purchase Stage Identification on Imbalanced Twitter Data”. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. 2017.
- [4] 井上結衣, 藤井敦. “質問応答サイトを用いた意見テキストの収集と極性反転文の検出”. 言語処理学会第 16 回年次大会発表論文集, pp.106-109, 2010.
- [5] Wen-Tau Yih, Joshua Goodman, and Vitor R. Carvalho. “Finding advertising keywords on web pages”. In *Proceedings of international conference on World Wide Web*, 2006.
- [6] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦. 日本語語彙大系 CDROM 版, 岩波書店, 1997.
- [7] 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, Vol.12, No.2, pp.203-222, 2005.
- [8] 東山昌彦, 乾健太郎, 松本裕治. “述語の選択好性に着目した名詞評価極性の獲得”. 言語処理学会第 14 回年次大会論文集, pp.584-587, 2008.