

# レビュー文書における重要文選択の一手法

松尾 哉太<sup>1,a)</sup> 新妻 弘崇<sup>1,b)</sup> 太田 学<sup>1,c)</sup>

**概要:** 通販サイト等の Web 上で公開されているレビューは、ユーザが商品の購入を検討する上で有益な情報源である。しかし、大量に投稿されたレビューを全て閲覧し、その内容を把握するのは困難である。そこで本研究では、レビュー閲覧の労力削減、内容把握の支援を目的として、レビュー文書中の重要文を自動で選択する手法を提案する。本研究では、商品に対する評価表現を含む文をレビュー文書における重要文と仮定し、複数の指標を用いて、レビュー文書中の商品に対する評価を含む文における指標値を算出する。その指標値を基に重要文を選択する。実験では、選択した重要文をあらかじめ作成した正解データと比較し、考察する。

## An Important Sentence Selection Method for Review Documents

KANATA MATSUO<sup>1,a)</sup> HIROTA KA NIITSUMA<sup>1,b)</sup> MANABU OHTA<sup>1,c)</sup>

### 1. はじめに

近年、Web 上の通販サイトを利用した商品取引が増加している。それに伴い、通販サイトに商品に関するユーザの意見や感想を含むレビューが多く投稿されている。例えば、そのような通販サイトの一つである楽天市場<sup>\*1</sup>では、「みんなのレビュー」というサービスがある(図1)。これを利用すると、商品レビューの投稿、閲覧が容易に行える。図1のようにレビューは商品に対する評価値(レビュー評価)を決定し、商品に関する意見や感想をレビューとして記述できる。図1からは、「液晶」や「使い勝手」といった商品属性について良い評価がされていること、なおかつレビュー評価が最高の“5”であることが見て取れる。このようなレビューを閲覧して得られる商品に対する意見や感想は、商品の購入を検討する際の重要な情報となる。

しかし大量に投稿されたレビューに関して、的確に内容を把握することは容易ではなく、多大な労力がかかる。そこで本研究では、レビュー閲覧の労力削減、内容把握を支援することを目的とし、レビュー文書から重要文を自動で

選択する手法を提案する。本研究における重要文とは、“液晶が非常に綺麗です。”のようにレビューが対象の商品に対しての評価や感想を述べている、または商品選択の際のポイントとなる点が記述されているような文を想定する。提案手法では、レビュー文書から商品の属性に関する評価が記述されている箇所を抽出し、それが含まれている一文に対して、商品の評価を表す語のスコアや文節数等の重要度を表す複数の指標を算出する。そして、その値を基にレビュー文書から重要文を選択する。

本稿では、2節で関連研究について述べ、3節でレビュー文書からの評価抽出手法を説明する。4節では評価属性が得られない場合に評価属性を推定する手法を説明する。つづく5節でレビュー文書中から重要文を選択する手法を説明し、6節で提案手法の評価実験を行う。7節でまとめと今後の課題について述べる。

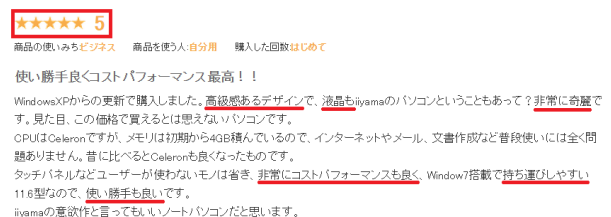


図1 商品に関するユーザのレビュー(楽天市場)

<sup>1</sup> 岡山大学大学院自然科学研究科

a) matsuo@de.cs.okayama-u.ac.jp

b) niitsuma@suri.cs.okayama-u.ac.jp

c) ohta@suri.cs.okayama-u.ac.jp

\*1 <http://www.rakuten.co.jp/>

## 2. 関連研究

商品レビューや商品に対するツイートから、商品の特徴を表す語と評価を示す語を抽出し、商品の評価の表示や要約を目的とする研究が数多く行われている [1][2][3]。これらの研究では、商品の特徴や属性を表す語と、その評価を表す語をペアで抽出し、それらにポジティブ、ネガティブといった極性、スコアを付与した結果から全体の評価を集約、提示する手法が数多く用いられている。

谷本ら [4] は、レビューサイトに投稿されたレビュー文書を用いて特定カテゴリの商品に関する評価表現辞書を自動生成し、商品属性ごとに評価を分かり易く可視化する手法を提案した。本研究におけるレビュー文書からの評価抽出は彼らの手法を参考にした。

柏木ら [5] は、レビュー文書中の省略された属性を推定し、意見情報を抽出する手法を提案した。彼らは評価を示す語、属性を示す語をアノテーションしたデータと、推定する属性の候補のリストを作成し、機械学習を用いて属性を推定した。本研究でも商品属性の推定を行なうが、彼らと異なり本稿では機械学習を用いず、商品属性クラスを手動で作成し、評価語の手動分類結果から商品属性を推定する。

平尾ら [6]、飯田ら [7] は機械学習や tf-idf 等を素性として用いて新聞記事から重要文を抽出する手法を提案した。文書から重要文を抽出、選択するという目的は本研究と同じだが、本研究では記述様式が定まっていないレビュー文書を対象とし、レビュー文書特有の表現である商品に対する評価や感想に着目して重要文選択を行うという点が彼らと異なる。

## 3. 評価抽出

レビュー文書には、主に商品の気に入った点や良い点、悪い点が記述されていることが多い。そこで本研究では、レビューが商品に対しての評価や感想を述べている文を重要文と仮定する。そのため、レビュー文書から商品の属性語と評価語を抽出する。本節では、その手法について説明する。なお本稿で属性語とは、「デザイン」や「値段」のような商品の属性を表す語（名詞）であり、評価語とは、「良い」や「安い」のような商品の評価を表す語（形容詞、形容動詞）とする。

### 3.1 商品属性と評価語

本研究では、対象とする商品ジャンルをあらかじめ定め、通販サイト価格.com<sup>\*2</sup> を参考に商品属性、属性語を定義する。また本稿では、対象とする商品ジャンルは「ノート PC、タブレット PC」とし、「デザイン」、「処理速度」、「液

晶」、「グラフィック」、「使いやすさ」、「携帯性」、「バッテリー」、「拡張性」、「価格」、「配送」、「総評」の 11 の商品属性を定義した。またそれぞれの商品属性を表していると考えられる、レビュー文書中に頻出する語を、あらかじめ属性語として複数定義した。定義した商品属性、属性語を表 1 に示す。また日本語 WordNet<sup>\*3</sup> を用いて、定義した属性語を拡充する。日本語 WordNet の上位語もしくは下位語にあらかじめ定義した属性語が該当する場合、その名詞は該当する属性語が分類される商品属性を表すものとする。

### 3.2 属性語と評価語の抽出

本稿では、以前に我々が提案した手法 [12] でレビュー文書から属性語と評価語を抽出する。レビュー文書からの属性語と評価語の抽出の手順を以下に示す。

- (1) レビュー文書集合から特定の商品ジャンルにおけるレビュー文書を収集する。
- (2) 得られたレビュー文書のレビュー本文を単位文に分割する。
- (3) 単位文を係り受け解析する。
- (4) 係り受け解析した結果、形容詞または形容動詞語幹の名詞と判定された語を評価語として抽出する。
- (5) 評価語と係り受け関係にある名詞がある場合、その名詞と評価語のペア（評価ペア）を抽出する。
- (6) 抽出した評価ペアの名詞の商品属性が 3.1 節の方法で決まらない、もしくは評価語と係り受け関係にある名詞がない場合は、その評価語（単独評価語）の商品属性を推定する。

ここで単位文とは、テキストにおける区切り文字で区切られた一文のことであり、評価抽出は単位文ごとに行う。また本研究で利用する区切り文字とは以下の記号のことである。

表 1 商品カテゴリ「ノート PC、タブレット PC」について定義した商品属性と属性語

商品属性	属性語
デザイン	デザイン, 色, 質感, 外装, 外観, 本体
処理速度	処理, 速度, 性能, 機能, スペック, パフォーマンス, スピード, 動作, 起動, CPU
グラフィック	グラフィック, 文字, 表示, 画質
拡張性	拡張, 添付, 付属, ソフト
使いやすさ	使い勝手, 操作, キーボード, 設定
携帯性	携帯, サイズ, 大きさ, 重さ, 持ち運び
バッテリー	バッテリー, バッテリー, 駆動
液晶	液晶, 画面, ディスプレイ
価格	価格, 値段, コスト, コストパフォーマンス
配送	配送, 到着, 発送

\*2 <https://kakaku.com/>

\*3 <http://compling.hss.ntu.edu.sg/wnja/>

- “.”, “。”, “?”, “?”, “!”, “!”, “【”

なお、区切り文字が連続する場合、それらの最後の文字を文の区切りとする。本研究では係り受け解析器の CaboCha<sup>\*4</sup> を用いて係り受け解析を行うことで、評価語と係り受け関係にある名詞の有無を判定する。(6)の商品属性推定は4節で説明する。

#### 4. 商品属性の推定

ここでは3.2節で説明した属性語と評価語のペアの抽出において、商品属性が属性語から判別できない評価語のペア (判別不能評価語)、また名詞との係り受け関係が得られない評価語 (単独評価語) について、評価語のみから商品属性を推定する手法を説明する。本研究では対象とするレビューデータとして、2012年1月に楽天市場に投稿されたレビューを用いる [13]。以後、このデータを楽天データと呼ぶ。ここでは、楽天データの商品カテゴリ「ノートPC, タブレットPC」のレビューから無作為に選んだ300件のレビューにおいて、10回以上出現する単独評価語と判別不能評価語の両方について商品属性を推定する。この目的のために、そのレビュー文書を読んでこれらの評価語を手であらかじめ定義した商品属性クラスに分類した。集計した評価語は名詞 (属性語) が得られなかった単独評価

表2 レビュー文書から抽出した評価語と出現回数 (カテゴリ: ノートPC, タブレットPC)

評価語	出現回数	評価語	出現回数	評価語	出現回数
良い	107	早い	28	小さい	11
満足	65	大きい	17	高い	11
ない	63	速い	17	便利	11
安い	50	悪い	16	軽い	10
十分	38	遅い	14	快適	10
綺麗	31	重い	14		
いい	30	うれしい	11		

表3 商品カテゴリ「ノートPC, タブレットPC」における評価語クラス分類結果の例

商品属性	評価語			
	良い	満足	速い	遅い
デザイン	11	4	0	0
処理速度	10	10	17	11
液晶	2	0	0	0
グラフィック	1	0	0	0
使いやすさ	6	1	0	0
携帯性	5	1	0	0
バッテリー	3	3	0	0
拡張性	2	0	0	0
価格	19	12	0	0
配送	1	0	0	2
総評	18	24	0	0
その他	29	10	0	1

\*4 <https://code.google.com/p/cabocha/>

語と、名詞とペアで出現した評価語の両方を含む。単独評価語は、全591件のうち214件 (36.2%) である。商品属性クラスは3.1節で述べた商品属性に「その他」を加えたものである。抽出した評価語と、その出現回数を数えた結果を表2に示す。また、評価語をレビューの文脈に応じて商品属性クラスに人手で分類した例を表3に示す。商品カテゴリ「ノートPC, タブレットPC」では、「良い」や「満足」といったユーザの感情を表す語に加え、「速い」や「遅い」のような処理速度を表す語が多く出現している。

本稿では、表2の結果を利用して、以下の2つの方法で判別不能評価語と単独評価語の商品属性を推定する。

- (1) 評価語を分類した商品属性クラスに基づき商品属性を推定する。
- (2) (1) で商品属性が推定できなかった評価語については、レビュー文書中でその評価語が出現した直前に述べられている商品属性と推定する。

#### 4.1 評価語を分類した商品属性クラスに基づく商品属性推定

表2に示した評価語の商品属性の決定方法を説明する。本稿では、以下の2種類の手法で商品属性を定め、それぞれの結果を比較した。

(手法1) 各評価語について、商品属性クラス毎の出現回数をその評価語の総出現回数で割った比  $AppR$  を算出し、閾値を超えたクラスを推定商品属性とする。

(手法2)  $AppR$  に商品属性クラスの出現確率を重みとして掛け合わせた値  $AppR_w$  を算出し、閾値を超えたクラスを推定商品属性とする。

評価語の推定商品属性は、なるべく少ないことが望ましい。そこで、本節の推定では、2つ以下の商品属性に絞ることが可能な評価語のみ商品属性を推定する。また、「その他」は推定する商品属性クラスに含めない。

#### 4.2 商品属性推定のための予備実験

本節では4.1節で説明した商品属性推定のために行った予備実験について説明する。4.1節の2つの手法による推定結果と、その推定結果を用いて評価語毎に算出した網羅率を表4, 表5に示す。ここで網羅率とは、調査した300件のレビューにおいて、各評価語から推定できた商品属性における評価語の出現数を、その評価語の総出現数で割った値のことである。

本実験では、評価語1語につき「その他」以外の商品属性を最大2つ推定した。また本稿では、このように商品属性推定できる評価語を最も多くするために、 $AppR$  の閾値を0.3,  $AppR_w$  の閾値を0.04として実験を行った。表4,

表5において、閾値に満たない商品属性は“(デザイン)”のようにかっこがきで示している。また網羅率“(0.554)”のようにかっこがきで示しているものは、かっこがきの商品属性を含めた網羅率のことである。また表4、表5の「速度」は「処理速度」のことである。

この結果、手法(1)が手法(2)に比べ、商品属性が推定可能な評価語が多く、各評価語における網羅率に大きな差は見られなかった。本稿ではより多くの評価語から商品属性を推定することを目的とするため、手法(1)を用いて商品属性推定を行う。

#### 4.3 直前の属性を用いた商品属性推定

本研究では、4.1節の方法で商品属性が推定できなかった評価語について、同文中で直前に述べられている属性を用いて商品属性を推定する。

図2に例を示す。「処理速度が速くて、私は良いと思います」という文を係り受け解析し評価抽出を行うと、「処理速度-速い」という評価ペアと、「私-良い」という評価ペアが抽出される。ここで「私-良い」の評価ペアに関して、「私」という名詞は属性語として定義されておらず、また「良い」という評価語からも4.1節の方法では属性推定ができない。これは表3に示したように、「良い」は様々な商品属性の評価に使用されるからである。しかしこの場合、直前の「処理速度-速い」の評価ペアから、「処理速度」という商品属性について評価していることが分かる。このことから、「私-良い」の評価ペアも「処理速度」についての評価とする。すなわち、同文中で直前に述べられている属性がある場合、その属性を評価語の推定商品属性とする。なお、本節で商品属性を推定する評価語は、4.2節のものを除く。

### 5. 重要文選択

本節ではレビュー文書から重要文を選択する手法について説明する。本稿における重要文とは、レビュー文書中で商品についての評価や感想が述べられている文と仮定している。また、重要文選択に用いる指標を3種類定義する。

表4 AppR の閾値 0.3 における商品属性推定結果

評価語	推定商品属性	網羅率	評価語	推定商品属性	網羅率
満足	総評, (価格)	0.369(0.554)	遅い	速度, (配送)	0.786(0.929)
安い	価格	1.00	重い	携帯性, (速度)	0.714(0.929)
十分	速度, (総評)	0.342(0.553)	小さい	携帯性, (液晶)	0.364(0.545)
綺麗	液晶, (デザイン)	0.355(0.451)	高い	速度, 価格	0.909
早い	速度, 配送	0.857	軽い	携帯性	0.700
大きい	携帯性, 液晶	0.941	快適	速度	0.400
速い	速度	1.00			

表5 AppR<sub>w</sub> の閾値 0.04 における商品属性推定結果

評価語	推定商品属性	網羅率	評価語	推定商品属性	網羅率
満足	総評, (速度)	0.369(0.523)	遅い	速度, (配送)	0.786(0.929)
安い	価格	1.00	重い	携帯性, 速度	0.714(0.929)
十分	速度, (総評)	0.342(0.553)	高い	速度, 価格	0.909
早い	速度, (配送)	0.429(0.857)	軽い	携帯性, (速度)	0.700(0.800)
速い	速度	1.00	快適	速度	0.400

これらの3種類の指標の詳細は5.1節で説明する。これらの指標の値(指標値)を3.2節の方法で抽出した全ての評価ペアまたは単独評価語ごとに算出する。そして、その評価ペアまたは単独評価語を含む単位文を重要文候補とし、重要文候補の中からこれらの指標値に基づき重要文を選択する。重要文選択の手法は5.2節で説明する。本研究では、レビュー閲覧の労力削減を目的とするため、レビュー文書1件につき最大3文まで重要文を選択すると定めた。

#### 5.1 重要文選択に用いる指標

本研究では、重要文選択に以下の3種類の指標を用いる。

- (1) 単位文中に出現する評価ペアと単独評価語の極性値: Polarity
- (2) 評価ペアと単独評価語における商品属性の  $t$  値:  $t$ -statistic
- (3) 単位文に占める評価語の割合: E-ratio

まず、3.2節の手法を用いてレビュー文書から抽出した全ての評価ペアと単独評価語ごとにこれらの指標値を算出する。単位文中に複数の評価ペアまたは単独評価語が出現する場合は、それらの指標値を合計したものをその単位文におけるそれぞれの指標値とする。その指標値に基づき選択する重要文を定める。

##### 5.1.1 単位文中に出現する評価ペアと単独評価語の極性値

抽出した評価ペアの評価語と単独評価語に対して、乾ら[8][9]が公開している日本語評価極性辞書を用いてスコアを付与する。このスコアを本稿では極性値 Polarity と呼ぶ。本研究では、この評価極性辞書にレビュー文書に頻出する形容詞、形容動詞語幹の名詞をさらに追加した。辞書に追加した単語とそのスコアを表6に示す。そのスコアに対し、我々が[12]で用いた手法で重みを掛け、極性値 Polarity を算出する。具体的には、「速くない」のように否定の助動詞「ない」が評価語と同時に出現する際は、「速い」のスコアの正負を逆転させる。すなわち重みとして-1

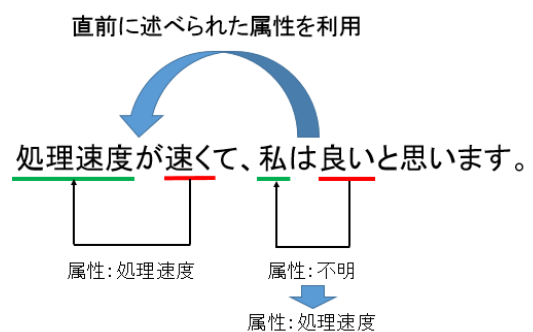


図2 直前の属性を用いた商品属性推定の例

表 6 日本語評価極性辞書に追加した単語とスコア

スコア	追加した単語
1.0	強い, 速い, 長い, 易い, 大きい, 格好いい, 綺麗, 見やすい
-1.0	小さい, 少ない, 高い, 短い, 弱い

表 7 強調語, 驚嘆表現

重み	強調語, 驚嘆表現
2.0	かなり, 極めて, 非常に, とても, 超, 随分, 大いに, 実に, 絶対, 特に, 最も, はっきり, めっちゃ, 「!」が連続する場合
1.5	少しも, 結構, 益々, 更に, 特別, なかなか, 普通に, やたら, むしろ, !, !, ○, ◎, ♪, ★, ☆, ^
0.8	若干, 少し, やや, ちょっと, まあ, わずかに, 少々, 何となく, 多少, まずまず, 意外と, とりあえず
0.5	決して, 大して, それほど, そんなに, あまり, あんまり

を掛ける。さらに、レビュー文書には「とても」や「あまり」のような評価語の働きに影響を与える語（強調語）が数多く出現する。また、「!」や「◎」のような記号（驚嘆表現）を用いて評価されていることもある。そこで係り受け解析の結果、強調語が評価語に係っている、また評価語と同じ文中に驚嘆表現が含まれる場合、評価語のスコアにあらかじめ定めた重みを掛け合わせる。定めた強調語、驚嘆表現とその重みを表 7 にまとめる。このように、スコアに重みを掛けたものが評価語の Polarity である。

例えば、「処理速度はあまり速くないです。」という単位文では、「速い」という評価語に「あまり」という強調語と「ない」という否定の助動詞に係っている。この場合、「速い」のスコア 1.0 に強調語「あまり」の重み 0.5 を掛け合わせ、その値の正負を逆転させた -0.5 をこの単位文の Polarity とする。Polarity の範囲は [-4.0, 4.0] である。Polarity が負の場合は商品に対する悪い評価、指標値が正の場合は商品に対する良い評価が述べられているとする。

### 5.1.2 評価ペアと単独評価語における商品属性の $t$ 値

ここでいう  $t$  値は、それぞれの属性のレビュー評価への影響の大きさを表す値で、重回帰分析を用いて算出することができる [12]。そのため  $t$  値を指標とすることで、レビュー評価の根拠となりうる商品属性を示すことを目的とする。この  $t$  値の算出にあたり、我々が作成した以下のような重回帰分析モデルを使用する。

まず、 $t$  値算出のため、5.1.1 節の方法で得られた極性値から、商品属性ごとの極性値を算出する。評価ペアの名詞が属性語の場合、その属性語が分類される商品属性に評価ペアの評価語の極性値を加える。評価ペアの名詞が属性語でない、または単独評価語の場合は商品属性推定を行い、

表 8 ノート PC, タブレット PC のレビューの重回帰分析における  $t$  値

商品属性	$t$ 値
デザイン	3.7787
処理速度	2.29508
グラフィック	1.1120
拡張性	0.87585
使いやすさ	2.5553
携帯性	0.088365
バッテリー	-0.53349
液晶	3.1958
価格	2.5013
配送	1.7774
総評	2.0290

推定できる場合は推定した商品属性に極性値を加える。推定できない場合は商品属性「その他」に極性値を加える。そして星の数などで表されるレビュー評価を目的変数、商品属性ごとの極性値を説明変数として重回帰分析を行い、レビュー評価の根拠となりうる商品属性を示す値として、 $t$  値を算出する。

本稿では、楽天データにおける商品カテゴリ「ノート PC, タブレット PC」のレビュー 100 件を用いてこの重回帰分析を行い、商品属性ごとの  $t$  値を算出した。なお、商品属性「その他」はレビュー評価の明確な根拠とならないと考え、その  $t$  値は指標値の対象としない。算出した  $t$  値を表 8 に示す。本研究では、評価に与える影響の大きさを表す指標として算出した  $t$  値を用いる。そのため、抽出した評価ペアまたは単独評価語で、得られた商品属性に該当する  $t$  値の絶対値を指標値  $t$ -statistic とする。例えば、「速度-速い」という評価ペアが得られた場合、このペアの  $t$ -statistic は、「処理速度」の  $t$  値である 2.29508 となる。

### 5.1.3 単位文における評価語の割合

レビュー文書において、レビューが最も主張したい話題が記述されている文では、評価語が多く用られ、また商品詳細に説明するため文が長くなる傾向にある。そこで、抽出した評価ペアが存在する単位文における評価語の割合を指標として用いる。具体的には、単位文中の評価語の数をその単位文の文節数で割ったものを指標値 E-ratio とする。

## 5.2 指標に基づく重要文選択

本研究では、5.1 節で説明した指標を用いて重要文を選択する。具体的には、以下の 2 つの手法を用いて重要文を選択する。

- (手法 1) それぞれの指標値を足し合わせた sum を用いて重要文を選択
- (手法 2) libSVM を用いて重要文を選択

手法1のsumを式1に示す。ここでPolarityとt-statisticは負の値をとり得るため、それぞれ絶対値を足し合わせる。

$$\text{sum} = |\text{Polarity}| + |\text{t-statistic}| + \text{E-ratio} \quad (1)$$

本研究では、評価ペアまたは単独評価語を含む単位文を重要文候補として、sumを算出する。そして重要文候補中からsumの上位3文を重要文として選択する。

手法2で用いるlibSVM<sup>\*5</sup>は、パターン識別器の代表的手法として広く用いられているサポートベクタマシン(SVM)のライブラリである。本研究では3つの指標値を3次元の特徴ベクトルとして、重要文、非重要文の特徴を学習し、レビュー文書中の各単位文を重要文、非重要文に分類する。

## 6. 評価実験

提案手法を用いてレビュー文書から重要文を選択する実験を行った。本節では5.2節の方法で重要文を選択した結果を比較し、選択した重要文について考察する。

### 6.1 重要文の選択実験

5.1節で述べた手法を用いて選択した重要文を、人が選んだ重要文である正解データと比較した。実験のため、2012年1月、2月に投稿された商品カテゴリ「ノートPC、タブレットPC」の楽天データのうち、4節のデータを一部含む100件を用いて正解データを作成した。ここで正解データとは、岡山大学工学部情報系学科の学生3名が各レビュー文書を閲覧し、2名以上が重要文と選択したものをそのレビューにおける重要文としたものである。

sumおよびlibSVMを用いて重要文を選択した結果と、正解データのレビュー文書における単位文を全て重要文とみなした結果をそれぞれ正解データと比較し、再現率、適合率、F値、要約率を算出した。また、5.1節で述べた3つの指標値の有効性を検証するため、それぞれの指標値のみを用いて、絶対値の大きいものから重要文として選択した結果を正解データと比較し、再現率、適合率、F値、要約率を算出した。その結果を表9に示す。ここで要約率とは、選択した重要文の数をレビュー文書中の全単位文の数で割ったものである。なお、libSVMは予習を伴うため5分割交差検定の結果を示している。

また、平尾ら[6]と飯田[7]らが行った重要文抽出の結果と比較するため、sum、libSVMの結果のF値と要約率を表10にまとめる。sumを用いて重要文を選択した結果のレビュー文書の要約率は約39%(229/580)、libSVMの結果の要約率は約40%(230/580)であった。そのため、平尾ら、飯田らの実験結果からそれに近い要約率30%、50%のものを選択した。なお、平尾らは新聞コーパスを用いた日本語要約評価用データセットであるTSC[10]のデータ

\*5 <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>

表9 重要文判定結果

手法	再現率	適合率	F値	要約率
Polarityのみ	0.558	0.677	0.611	39%
t-statisticのみ	0.514	<b>0.722</b>	0.601	34%
E-ratioのみ	0.561	0.681	0.615	39%
sum	0.572	0.694	0.627	39%
libSVM	0.625	0.701	<b>0.661</b>	40%
レビュー文書中の全単位文	<b>1.00</b>	0.479	0.648	100%

表10 平尾ら、飯田らの実験結果との比較

手法	F値	要約率
平尾ら	0.483	30%
飯田ら	0.463	30%
平尾ら	0.628	50%
飯田ら	0.626	50%
sum	0.627	39%
libSVM	0.661	40%

を、飯田らはTSC2[11]のデータを利用している。よって実験データの性質が異なるため、表10の比較は参考のためのものである。

表9から、本手法を用いた重要文選択の結果、libSVMのF値が最も高くなり、単位文を全て重要文とみなした結果のF値を上回った。それぞれの指標値のみを用いた結果の中では、E-ratioのF値がPolarity、t-statisticのF値より高くなった。適合率では、本手法で選択した重要文の結果が単位文を全て重要文とした結果を上回った。ここで、本手法では重要文を最大3文しか選択しないため、結果的に再現率よりも適合率を重視する選択となった。また要約率が小さいと、再現率も低くなる。

表10から、既存研究との比較では、sumを用いて選択した重要文の要約率が39%となり、平尾ら、飯田らの30%の結果と比較すると、要約率は高いが、F値も彼らの結果より高かった。平尾ら、飯田らの50%の結果と比較すると、要約率は低い、F値はほぼ同じ結果となった。同様に、libSVMを用いて選択した重要文の要約率が40%となり、平尾ら、飯田らの30%の結果と比較すると、要約率は高いが、F値も彼らの結果より高かった。平尾ら、飯田らの50%の結果と比較すると、要約率は低い、F値は彼らの結果より高かった。

これらの結果を踏まえ、6.2節で実験データにおける重要文を分析し、精度向上のための考察を行う。

### 6.2 選択した重要文の考察

6.1節で用いた実験データの内訳を表11に示す。表11から、レビュー文書100件の全580文のうち、重要文は278文ある。そのうち、本手法を用いて抽出できなかった重要文は、評価語を含むが抽出できなかったものが25文、評価語を含まないものが61文あった。表12にそれぞれの例を示す。

表 11 実験データの内訳

重要文		非重要文	
評価語を含む	評価語を含まない	評価語を含む	評価語を含まない
抽出	非抽出	61	103
191	25		

表 12 選択できなかった重要文の例

評価語を含む重要文	評価語を含まない重要文
自宅でしか使っていないですけど、良いですよ^^	ネット、動画再生などはサクサク動きます。
Wifi すぐつながり、快適。	操作性は劣って見えてしまう。
【処理速度】9年ぶりに新調したので、大満足です	マウスの使い方が似ていて、しっくりきます。

表 12 に示したように、評価語を含むが抽出できなかった重要文には、「自宅でしか使っていないですけど、良いですよ^^」といった区切り文字が使われていない文があった。区切り文字が使われない文を係り受け解析すると解析を誤ることが多いため、適切な前処理を行いレビューの文章を正しい単位文に分割する必要がある。また、「満足」や「快適」といった語が名詞として判定され、評価語として抽出できないこともあった。評価語を含まない重要文の中には、「サクサク動く」や「劣って見える」のように商品の評価を表す語として名詞や動詞等の語が用いられているものがあつた。このような語は本手法では評価語と扱わないため、重要文候補として抽出することができなかった。そこで、名詞や動詞を評価語として追加することを検討している。

一方、評価語を含む非重要文では、「大変良い商品でした」や「個人的には満足です」のように商品の総評やレビューの感想として評価語が使われている文が多かった。この結果から、抽出した評価語の評価語が商品の総評として使われる語かつ商品属性が推定できない場合は、重要文として選択しないことを検討している。

## 7. まとめと今後の課題

本稿では、レビュー閲覧の内容把握支援と負担軽減のために、レビュー文書における重要文を自動で選択する方法を提案した。提案手法では、まずレビュー文書から評価ペアもしくは単独評価語を抽出し、それらを含む単位文を重要文候補とする。そして重要文候補の中から、評価ペアの評価語または単独評価語の極性値、商品属性の  $t$  値、単位文における評価語の割合の 3 つの指標値を基に重要文を選択する。

評価実験では、レビュー文書中の全ての単位文を重要文とみなした結果は再現率 1.00、適合率 0.480、F 値 0.648 となり、指標値の絶対値を足し合わせた sum を用いて重要文を選択した結果は再現率 0.572、適合率 0.694、F 値 0.627、libSVM を用いて重要文を選択した結果は再現率 0.625、適合率 0.701、F 値 0.661 となった。単位文を全て重要文と

みなした結果 (要約率 100%) と比較して、適合率は改善し、libSVM の F 値が上回つた。

提案手法で選択できた重要文と選択できなかった重要文について調べた結果、正しく単位文に区切ることができない文は係り受け解析が失敗し、評価語が抽出できないという問題があつたため、前処理で正しく単位文に分割することを検討している。今後の課題には、抽出できる評価語を増やし、選択できなかった評価語を含む重要文を選択できるようにすることが挙げられる。また、他の商品カテゴリにおけるレビュー文書を対象とした重要文選択実験を行うことを検討している。

**謝辞** 本研究を進めるにあたり、楽天データ公開において公開された楽天市場「みんなのレビュー・口コミ情報」を使用させて頂きました。データを公開して頂きました楽天株式会社に深く感謝致します。

## 参考文献

- [1] 平山拓央, 湯元高行, 新居学, 高橋豊, “属性評価モデルに基づく商品評価の抽出と提示”, DEIM Forum 2011, F2-5, 2011.
- [2] 駒田康孝, 山名早人, “商品評価ツイートからの属性語自動抽出手法の提案”, DEIM Forum 2014, B5-6, 2014.
- [3] 中野裕介, 湯本高行, 新居学, 佐藤邦弘, “商品レビュー要約のための属性-意見ペア抽出”, 情報処理学会研究報告, DBS-160, No.15, 2014.
- [4] 谷本融紀, 太田学, “評価属性を考慮した評判情報の可視化”, 情報処理学会研究報告, DBS-151, No.12, 2010.
- [5] 柏木潔, 小町守, 松本裕治, “レビュー文書からの省略された属性の推定を含めた意見情報抽出”, 言語処理学会第 19 回年次大会, B5-5, 2013.
- [6] 平尾努, 磯崎秀樹, 前田英作, 松本裕治, “Support Vector Machine を用いた重要文抽出法”, 情報処理学会論文誌, Vol.44, No.8, pp.2230-2243, 2003.
- [7] 飯田龍, 徳永健伸, “談話の顕現性を考慮した重要語抽出とその応用”, 研究報告自然言語処理 (NL), Vol.2009-NL-193, No.9, 2009.
- [8] 公開資源/日本語評価極性辞書-東北大学 乾・岡崎研究室, <http://www.cl.ecei.tohoku.ac.jp/index.php?Open%20Resources%2FJapanese%20Sentiment%20Polarity%20Dictionary>
- [9] 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.3, pp.203-222, 2005.
- [10] Fukushima, T., Okumura, M., “Text Summarization Challenge”, Text Summarization Evaluation in Japan, Proc. NAACL2001 Work-shop on Automatic Summarization, pp.51-59, 2001.
- [11] 奥村学, 福島孝博, “TSC2(Text Summarization Challenge 2) の目指すもの”, 情報処理学会情報学 (基礎研究会報告), FI-63-5, pp.33-39, 2001.
- [12] 松尾哉太, 新妻弘崇, 太田学, “レビュー解析に基づくユーザ評価の根拠提示の一手法”, 情報処理学会研究報告, DBS-160, No.14, 2014.
- [13] 楽天技術研究所, 楽天データ公開, <http://rit.rakuten.co.jp/rdr/index.html>, (accessed 2014-06-19).