

商品説明ページを用いた評価視点別評判情報提示システム

瀬藤 亮[†] 佐藤 哲司^{††}

[†] 筑波大学図書館情報専門学群 〒305-8550 茨城県つくば市春日 1-2

^{††} 筑波大学大学院図書館情報メディア研究科 〒305-8550 茨城県つくば市春日 1-2

E-mail: ^{††}satoh@slis.tsukuba.ac.jp

あらまし Web 上のクチコミサイトに蓄積された評判情報は、商品購入やマーケティングの参考として用いることができる。しかし、評判情報が膨大である場合、全てに目を通すには多大な負担がかかる。本論文では、評判情報の閲覧支援を目的として、評判情報を言及している対象物の側面（評価視点）に基づいて分類し提示するシステムを提案する。例えば、ホテルを対象とした評判情報では「部屋」、「風呂」、「レストラン」などの評価視点に評判情報が分類される。本システムでは、多様な対象物に対応するために、クチコミサイトでそれぞれの対象物について個別に用意された商品説明ページを用いて評価視点を評判情報記事から抽出する。本システムの有効性をホテル予約サイトを対象とした実験により評価する。

キーワード 評判情報, Web 情報システム, 分類

Feature-Based Reputation Browser Using Product Descriptions

Ryo SETO[†] and Tetsuji SATOH^{††}

[†] School of Library and Information Science University of Tsukuba

1-2, Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

^{††} Graduate School of Library, Information and Media Studies, University of Tsukuba

1-2, Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

E-mail: ^{††}satoh@slis.tsukuba.ac.jp

Abstract Online reputation articles in e-commerce sites are useful for product choice, marketing research, and so on. However, if the number of the articles is large, it is hard to read all. In this paper, we propose a system that helps users to read the articles by classifying articles based on product features that are referred to. In our system, the product features are extracted from product description pages in e-commerce sites. We experiment with the system in a hotel reservation site to show the effectiveness of our system.

Key words Reputation, Web Information System, Classification

1. はじめに

近年、インターネットの普及により、誰でも気軽に情報を発信することができる環境が整ってきた。個人が発信する情報の中には、特定の商品・サービス（以下、対象物）に関する評判情報も含まれている。特に、「Amazon.co.jp」^(注1)、「Yahoo!映画」^(注2)、「楽天トラベル」^(注3)などのクチコミサイトには、様々な対象物に関する評判情報が大量に蓄積されている。ある対象物についての評判情報は、その対象物に関わる企業や消費者が参考

に用いることがある。例えば、企業は自社の提供する対象物に関して消費者の感想を評判情報から知ることができる。また、消費者は対象物を購入する際に、あるいは購入した対象物に起きた不具合に対処する際に参考することができる。しかし、対象物によっては投稿された評判情報の記事数は数百件を超える場合がある。評判情報が膨大である場合、全てに目を通すには多大な負担がかかる。

本研究では、評判情報の閲覧を支援するシステムを提案する。本システムは、ユーザの嗜好や状況により関心のある「評価視点」が異なることを踏まえ、評判情報を「評価視点」に基づいて分類し提示する。ここで「評価視点」とは評判情報で言及されている対象物に関する様々な側面である。例えば、対象物がホテルである評判情報では「部屋」、「レストラン」、「風呂」な

(注1): <http://www.amazon.co.jp/>

(注2): <http://movies.yahoo.co.jp/>

(注3): <http://travel.rakuten.co.jp/>

どが評価視点になり得る．本システムを用いることで，ユーザは関心のある評価視点に関する評判情報のみを選択的に閲覧することができる．

多様な対象物に対応するためには，対象物が持つ評価視点をあらかじめ用意しておくことは現実的ではない．そこで，本システムでは評価視点をシステムに入力された評判情報から抽出することとし，抽出した評価視点をクチコミサイトにおいて対象物ごとの用意された商品説明ページを用いて洗練し，体系的に提示する．以下，本論文では，評判情報からの評価視点の抽出と説明ページを用いた評価視点の洗練をあわせて「評価視点の抽出」と呼ぶことにする．

以下，2. で関連研究について述べ，本研究の位置づけを明確にする．3. で商品説明ページを用いた評価視点の抽出について述べ，4. で本システムについて述べる．5. で評価実験を行い，6. で結果を考察する．7. で本論文を総括する．

2. 先行研究と本研究の位置づけ

2.1 評判情報の閲覧を支援する先行研究の概要

評判情報の閲覧支援を目的とした先行研究では，評判情報を好評・不評に分類する場合，評判情報を評価視点に基づいて分類する場合がある．

評判情報の好評・不評による分類について，Pang ら [1]，Gamon ら [2] は，機会学習を用いる手法を提案している．藤村ら [3]，立石ら [4]，Hu ら [5] は「広い」「大きい」「良い」「満足した」などのように対象物あるいは評価視点の性質や，個人の主観的な感情を表す表現（評価表現）に着目する手法を提案している．

評判情報の評価視点に基いた分類について，立石らは分類項目として用いる評価視点をあらかじめ辞書として用意する手法を提案している．飯田ら [6] は機会学習を用いることで立石らの手法について精度を向上している．土田ら [7] は，立石らの手法を，クチコミサイトだけでなくブログから評判情報を収集できるよう拡張している．小林ら [8]，阿辺川ら [9] は，立石らの手法で用いる評価視点の辞書を効率的に構築する手法を提案している．Hu らは，評価視点の辞書を用いず，入力された評判情報そのものから評価視点を抽出している．具体的には，相関ルールマイニングを用いて出現頻度の高い語を評価視点として抽出している．同様に評価視点の辞書を用いない手法として，Scaffidi ら [10] は，一般的な文書と比較して入力された評判情報に出現しやすい語を評価視点として抽出している．Gamon らは，出現頻度と TF-IDF を複合的に用いて評価視点を抽出している．

2.2 本研究の位置づけ

本研究では，ユーザが評判情報を評価視点に基づいて選択的に閲覧することができるシステムを提案する．よって，評価視点に基づき分類する手法に焦点をあてる．Hu ら，Gamon ら，Scaffidi らの手法と同様に，入力された評判情報から評価視点を抽出する．評価視点の抽出では，評判情報中での語の出現状況だけでなく商品説明ページを用いることで精度を向上する．これらの特徴により，提案するシステムでは，事前に評価視点



図1 「ホテルステノ」の商品説明ページ (楽天トラベル)

の辞書を作成しておく必要がなく高い精度で評価視点を抽出することができる．

3. 商品説明ページを用いた評価視点の抽出

本章では，本システムの特徴である商品説明ページを用いた評価視点の抽出について述べる．まず，3.1 で商品説明ページの概要と，入出力を含めた全体の処理を概観する．そして，3.2～3.4 で個別の処理について述べる．

3.1 評価視点の抽出の概要

商品説明ページ

商品説明ページは，対象物である商品やサービスに関して企業等が用意した，いわゆる宣伝ページである．商品説明ページは，多くのクチコミサイトにおいて評判情報とともに提示されている．商品説明ページには，対象物の様々な側面が網羅的に記述されている場合が多い．一例として，楽天トラベルに登録されているホテル「ホテルステノ」の商品説明ページ^(注4)を図1に示す．「ホテルステノ」について「部屋」や「レストラン」といった構成要素をはじめとして「無線 LAN」が使えることや，周辺で「ゴルフ」ができることなど多様な側面が記述されていることが分かる．

このように対象物の多様な側面について記載された商品説明ページは，楽天トラベルに限られることなく「Amazon.co.jp」，「Yahoo!映画」など多くのクチコミサイトで提供されている．

入出力と処理の流れ

本システムの評価視点の抽出手法について，入力から出力までの処理の流れを図2に示す．入力は，ある対象物についての評判情報記事集合と商品説明ページである．まず，入力された評判情報記事から品詞情報を用いて評価視点になりやすい語（以下，評価視点候補）を抽出する．次に，評価視点候補に対して評判情報記事中における出現状況を用いて評価視点である可能性の高さによりスコアを付与する．そして，評価視点候補と商品説明ページに出現する語を照合し，商品説明ページに出現する評価視点候補のみ抽出する．以上で述べた3つの処理を段

(注4): <http://travel.rakuten.co.jp/HOTEL/2736/2736.html>



図2 入力から出力までの処理の流れ



図4 複数の接続ボタンが考えられる場合の抽出例



図3 複数名詞の抽出例

階的に行うことで、評判情報と商品説明ページから評価視点のリスト（以下、評価視点リスト）を生成する。リストされた評価視点は、スコアにより順位付けされている。

3.2 評価視点候補の抽出

評判情報記事から抽出する評価視点候補は、対象物の側面であることから名詞（複合名詞を含む）とした。まず、形態素解析器 MeCab^(注5)に評判情報記事を入力し、出力結果から名詞だけを抽出する。ただし、経験的に「名詞、非自立」、「名詞、代名詞」、「名詞、ナイ形容動詞語幹」、「名詞、特殊、助動詞語幹」、「名詞、接尾、助動詞語幹」、「名詞、接尾、人名」、「名詞、副詞可能」を品詞情報に持つ名詞は評価視点になりたくないため抽出しない。

また、複数の名詞が2回以上同じパターンで接続する場合、その名詞群を単一の複合名詞として抽出する。複合名詞の抽出例を図3に示す。同図左の記事を形態素解析すると、名詞「部屋」、「駐車」、「場」、「ホテル」、「総合」、「的」が得られる。「駐車」と「場」は2回接続するため「駐車場」として抽出される。「総合」と「的」は1回しか接続しないため複合名詞としては抽出されない。「駐車」は3回、「場」は2回出現しているが、そのうち2回は複合名詞「駐車場」として処理されるため「駐車」は2回、「場」は0回、「駐車場」は2回抽出される。

また3つ以上の名詞が接続し、複数の接続ボタンが考えられる場合がある。その場合は、最も出現頻度が多い接続ボタンから優先して抽出する。複数の接続ボタンが考えられる場合の抽出例を図4に示す。同図左の6件の記事を形態素解析すると名詞「東京」、「駅」、「周辺」が得られる。6件の記事のうち、「東京」と「駅」は3回、「駅」と「周辺」は2回接続している。そこで、出現頻度の多い「東京」と「駅」の組を「駅」と「周辺」の組よりも優先する。すなわち、「東京駅周辺」という文字列からは複合名詞「東京駅」と名詞「周辺」を抽出する。

接続ボタンの構成名詞数が3つ以上の場合も同様の方法で抽出の優先順位を決定する。接続ボタンの出現頻度が等しい場合は、構成名詞数が多い接続ボタンを優先する。この方法では、接続ボタンの出現頻度も構成名詞数も等しい場合はどちらを抽出するか決定できないが、ここでは先に出現した接続ボタンを優先することとした。

3.3 記事中の出現状況によるスコアリング

前節で抽出した評価視点候補に対して、評判情報記事での出現状況を用いて評価視点である可能性の高さに基づきスコアを付与する。スコアリングには、出現頻度、評価表現との特定パターンでの共起頻度（以下、共起ボタン）、評価表現との同一文内での共起頻度（以下、文内共起）の3種類の指標が考えられる。それぞれのスコアリング指標は独立であり、指標の優劣について後述の5.で評価する。

なお、2.で述べた先行研究では、出現頻度を基本とした複雑な指標を用いて評価視点を抽出する研究が多い。しかし、本研究では、次節で述べる商品説明ページを用いて評価視点を洗練することから、評価視点候補の評判情報記事中での出現頻度をそのままスコアとする簡便な手法を用いることとした。

共起ボタン

2.で述べたように、「評価表現」とは対象物や評価視点の性質、書き手の主観的な感情を表す表現である。例えば「広い」、「満足した」などである。立石ら[4]、小林ら[8]は評価表現と共起しやすい構文パターンを用いて評価視点を抽出している。この手法に基づき、評価視点候補について評判情報として現れやすい以下の2つの構文パターンで評価表現と共起した頻度をスコアとする。

(1) (評価視点候補)[が/は/も/に/を](評価表現)

例:「部屋 が 広い」

(2) (評価表現)(評価視点候補)

例:「広い 部屋」

なお、「部屋 がとても 広い」のように評価視点候補と評価表現が隣接していない場合を考え、共起頻度の計算はCabocha^(注6)により係り受け解析を行った文について行う。記事中の評価表現の同定は小林らにより公開されている評価表現辞書^(注7)を用いて行う。

(注5): <http://mecab.sourceforge.net/>

(注6): <http://chasen.org/taku/software/cabocha/>

(注7): http://www.syncha.org/evaluative_expressions.html

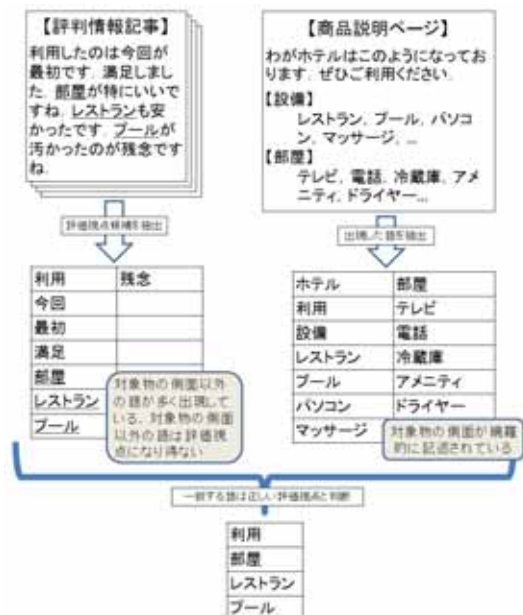


図 5 評価視点候補と商品説明ページに出現する語の照合例

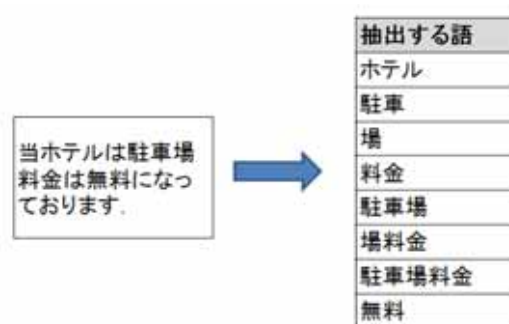


図 6 商品説明ページからの名詞・複合名詞の抽出例

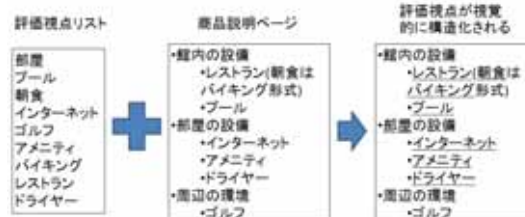


図 7 商品説明ページを用いて評価視点を視覚的に構造化する例

文内共起

評価視点として評価表現と近接して共起しているものの、上述の共起パターンには当てはまらない場合がある。そこで、評価表現と同一文内で共起した頻度をスコアとする手法を考える。文の区切りは「。」「，」「？」「！」「！」「！」「！」「！」「！」のいずれかの文字が1つ以上接続して出現する場合、あるいは、改行が2つ以上接続して出現する場合とする。

3.4 商品説明ページに出現する語との照合

商品説明ページには対象物の側面が網羅的に記述されている。一方、評判情報記事から抽出した評価視点候補には、対象物の側面ではない、評価視点になり得ない語が多く含まれている。

そこで、対象物の側面でない評価視点候補を除去するため商品説明ページに出現する語と照合する。照合例を図5に示す。「今回」「最初」など、先述の評価視点候補の抽出で誤って抽出された候補の多くを除去することができる。

ただし、商品説明ページに出現する語と一致する評価視点候補でも、評価視点ではない場合がある。例えば、図5の商品説明ページに書かれた「利用」という語は評判情報にも出現することから、誤って抽出される。そこで、3.3節で付与したスコアに対して閾値を設定し、閾値を下回る評価視点候補は抽出しない。

商品説明ページからの語の抽出方法について説明する。まず、HTMLソースから該当文字列を抜き出す。該当文字列とは、< BODY > < /BODY > タグに囲まれた部分のうち、以下の2つのいずれかに該当する文字列である。

(1) タグ以外の文字列

例「< a href="X" > トップ < /a > へ戻る」
「トップ」「へ戻る」

(2) 画像表示タグ中の代替文字列を表す属性altの属性値

例「< img alt="部屋の画像" src="X" >」

「部屋の画像」

抜き出した文字列を MeCab により形態素解析し、出力結果から名詞のみを抽出する。そして、名詞の接続処理を行う。ここでは接続パタンの出現頻度は考えず、全ての接続パタンの名詞群を抽出する。商品説明ページからの語の抽出例を図6に示す。同図左の商品説明ページの記述から、同図右の名詞・複合名詞を抽出する。

4. 提案システム

4.1 システムの概要

本システムは、評判情報の閲覧を支援することを目的として、評判情報を評価視点に基づき分類し提示する。本システムを用いることで、ユーザは評判情報を評価視点に基づき選択的に閲覧することができる。実装には Java, JSP を用いた。入力された評判情報記事集合と商品説明ページについて3.で述べた処理を行い、Webブラウザ上に出力する。

4.2 システムの機能

システムは、まずユーザに評価視点リストを提示する。ユーザが評価視点リストから評価視点を選択すると、その評価視点に言及した評判情報を提示する。

評価視点リストは、評価視点をただ羅列するのではなく視覚的に構造化し提示するのが望ましい。そこで、商品説明ページは人手で作成されているため可読性が高いことから、これを用いて評価視点を提示する。商品説明ページを用いて評価視点を視覚的に構造化した例を図7に示す。評価視点リストに商品説明ページを組み合わせることで「レストラン」と「朝食」など関連の強い語が近接して提示されている。

一方、3.で抽出しなかった評価視点候補の中にも、評価視点として適切な語が含まれている可能性がある。特に、3.で付与したスコアが極めて高い評価視点候補は、商品説明ページに出

現しているか否かに関わらず評価視点としてふさわしい場合が多い。これは企業が見落とした評価視点である可能性が高く、このような評価視点を提示することは消費者と企業双方にとって有用であると考えられる。以上を踏まえ、本システムではこれらの評価視点候補をスコアが高い順に整列し提示する。

評判情報の提示は文単位で行う。出現頻度でスコアリングする場合は、ユーザが選択した評価視点が出現した文を全て提示する。共起ボタンでスコアリングする場合は、ユーザが選択した評価視点が先述した構文ボタンで評価表現と共起した文のみを提示する。文内共起でスコアリングする場合は、ユーザが選択した評価視点が同一文内で評価表現と共起している文のみを提示する。また、文だけでは評価の全容が分からない場合を考慮し、その文が出現する記事の全文を提示することができる。文あるいは記事中のどこで評価視点に言及したか分かりやすくするために、評判情報中の評価視点は赤色で表示する。

4.3 システムの実行情例

システムの実行画面を図8に示す。同図は、楽天トラベルに登録されているホテル「ホテルステノ」に関する評判情報記事集合と商品説明ページを入力とした場合の例である。左部は、商品説明ページの該当文字列にリンクタグを埋め込み、評価視点を構造化して提示している。評価視点をマウスでクリックすると、画面右側の中段に評判情報が文単位で提示される。文単位の評判情報のいずれかをクリックすると、画面右側の下段にその文が出現した記事全文が提示される。同図は、ユーザが評価視点リストから評価視点「レストラン」を選択した状態である。レストランに言及した評判情報が文単位で提示されている。また、「レストランのお料理がとても気に入っていて、今回も楽しめました！」という文がクリックされ、この文が出現する記事の全文が提示されている。

商品説明ページに出現していない評価視点の閲覧は、右上段の「他の評価視点」というアンカーテキストをクリックすることで行う。実行例を図9に示す。商品説明ページに出現していない評価視点が、スコアにより順位付けされ画面中央に提示されている。

5. 評価

5.1 評価の概要

提案したシステムの構成要素について評価実験を行う。具体的には、(1) ユーザに提示する評価視点リスト中の評価視点が適切であるかどうか、(2) ユーザが選択した評価視点に関して提示した評判情報が妥当であるかどうかの2項目を評価する。

評価データとして用いる評判情報記事と商品説明ページは、楽天トラベルから4軒のホテルについて取得した。評価データについて、ホテル名、評判情報記事数、後述する正解ラベルが付与された評価視点数を表1に示す。評価視点数は重複を除き集計した。それぞれのホテルの商品説明ページは、楽天トラベルに登録された全てのホテルに対してそれぞれ用意された「詳細情報」ページを用いた。

提示した評価視点あるいは評判情報が正解であるかどうかを判断するために、事前に人手で評判情報中の評価視点と思われ

表1 評価データの詳細

ID	ホテル名	記事数	評価視点数
A	オークラフロンティアホテルつくば	79	122
B	ジャスマックプラザホテル	82	106
C	ホテルステノ	398	149
D	弓ヶ浜温泉 ホテル山海	291	115

る文字列に正解ラベルを付与した。正解ラベルの付与作業は筆者1人により行った。以下に例を示す。

- (a) 部屋と風呂が良かったです。
 - (b) レストランが開くのが遅く、部屋で待たされました。
- (a) について「部屋」と「風呂」に、(b) について「レストラン」に正解ラベルを付与する。「部屋」は(b)にも出現しているが、(b)は「部屋」に関する評判情報ではないため正解ラベルを付与しない。以上により、ここでは正解の評価視点は「部屋」「風呂」「レストラン」である。「部屋」に関する正解の評判情報は(a)、「風呂」に関する正解の評判情報は(a)、「レストラン」に関する正解の評判情報は(b)である。

評判情報記事数が100件を超えるホテルでは、正解ラベルの付与は最新記事から100件目までの記事に限り行った。この場合は、3.で述べた評価視点候補のスコアリングは全記事を用いて行い、100件目以内の記事に出現していない評価視点および文を評価の対象外とした。

5.2 提示する評価視点リストの評価

本システムにおける評価視点リストは2種類ある。商品説明ページに出現する評価視点のリストと、商品説明ページに出現しない評価視点のリストである。2種類の評価視点リストについて個別に評価を行う。

まず、商品説明ページに出現する評価視点のリストに関する評価実験について報告する。この評価視点リストは、3.で述べたように、正解の評価視点多く含まれている可能性が高い。ただし、商品説明ページに出現する語が、評判情報中で出現する評価視点全てをカバーしているとは限らない。これらを踏まえ、正確性と網羅性を評価するため、適合率と再現率を算出する。算出は、以下の式による。

$$\text{適合率} = \frac{\text{評価視点リスト中の正解の評価視点数}}{\text{評価視点リスト中の評価視点数}}$$

$$\text{再現率} = \frac{\text{評価視点リスト中の正解の評価視点数}}{\text{正解の評価視点数}}$$

商品説明ページに出現する評価視点を全て出力した場合は適合率が非常に低い値になるため、3.において付与したスコアに対して閾値を設定し、リストする評価視点を制限する。ここでは、閾値は経験的に以下の式により決定する。

$$\text{閾値} = \text{全評価視点候補中の最高スコア} \times 0.05$$

閾値は、3.で述べた3種類のスコアリング指標に対してそれぞれ独立に計算した。

商品説明ページを用いたことが評価視点の抽出に与える影響を確認するために、閾値を超えるスコアを持つ評価視点候補を全てリストした場合と比較した。それぞれの適合率と再現率を



図 8 ホテルステノの評判情報を評価視点別に提示した例



図 9 商品説明ページに出現しない評価視点を提示した例

表 2 に示す。

適合率は、全ホテルの全順位付け指標において大きく向上している。「他」、「感じ」、「お願い」といった、評判情報記事中に頻出する語を除去することができている。抽出ミスは「利用」、「確認」、「到着」など商品説明ページに含まれやすい語だった。一方、再現率は大幅に低下している。商品説明ページにより

誤って除去された評価視点は、抽出された評価視点の同義語、表記ゆれである語、側面を表す語が多かった。例えば、抽出されなかった評価視点「アメニティー」は、抽出された評価視点「アメニティ」の表記ゆれである。また、抽出されなかった評価視点「ランチ」は、抽出された評価視点「レストラン」の側面を表す。

表 2 商品説明ページを用いる場合と用いない場合についての評価視点の抽出の適合率と再現率

ID	結果	商品説明ページを用いる場合			商品説明ページを用いない場合		
		出現頻度	共起ボタン	文内共起	出現頻度	共起ボタン	文内共起
A	適合率	0.44	0.58	0.53	0.31	0.32	0.43
	再現率	0.11	0.06	0.14	0.20	0.10	0.25
B	適合率	0.50	0.47	0.50	0.31	0.28	0.27
	再現率	0.22	0.17	0.26	0.42	0.39	0.47
C	適合率	0.61	0.76	0.63	0.41	0.58	0.51
	再現率	0.15	0.13	0.17	0.26	0.22	0.31
D	適合率	0.45	0.80	0.50	0.27	0.51	0.37
	再現率	0.15	0.14	0.15	0.38	0.28	0.38
平均	適合率	0.50	0.65	0.54	0.33	0.42	0.40
	再現率	0.16	0.13	0.18	0.32	0.25	0.35

表 3 商品説明ページに出現しない評価視点の抽出の適合率と再現率

ID	結果	出現頻度	共起ボタン	文内共起
A	適合率	0.23	0.33	0.38
	再現率	0.03	0.02	0.03
B	適合率	0.32	0.11	0.40
	再現率	0.13	0.03	0.20
C	適合率	0.30	0.62	0.50
	再現率	0.06	0.08	0.08
D	適合率	0.24	0.43	0.33
	再現率	0.13	0.12	0.09
平均	適合率	0.27	0.37	0.40
	再現率	0.09	0.06	0.10

スコアリング手法別に見ると、出現頻度によるスコアリングと比較して、共起ボタンによるスコアリングでは適合率が大幅に向上している。文内共起によるスコアリングでは、適合率と再現率をともに高くできている。

次に、商品説明ページに出現しない評価視点のリストに関する評価実験について報告する。このリストは、先のリストに含まれていない評価視点を正確かつ網羅的に提示することを目的にしていることから、以下に示す適合率と再現率を用いて評価することとした。算出は以下の式による。

$$\text{適合率} = \frac{\text{評価視点リスト中の正解の評価視点数}}{\text{評価視点リスト中の評価視点数}}$$

$$\text{再現率} = \frac{\text{評価視点リスト中の正解の評価視点数}}{\text{商品説明ページに出現しない正解の評価視点数}}$$

商品説明ページに出現しない評価視点である点、評価視点をスコアに基づき整列して提示している点を踏まえ、閾値は先のリストに対する設定より高くする。ここでは、経験的に以下の式により決定する。

$$\text{閾値} = \text{全評価視点候補中の最高スコア} * 0.10$$

評価結果は、表 3 に示す通り、全体的に、適合率・再現率ともに低い値であった。システムが提示した評価視点を個々に調べたところ、商品説明ページに出現する評価視点の同義語や表記ゆれである語が多く、商品説明ページに出現していない概念を表す評価視点もいくつか確認することができた。

表 4 提示される評判情報の正解率

ID	出現頻度	共起ボタン	文内共起
A	0.75	0.83	0.79
B	0.70	0.74	0.73
C	0.68	0.77	0.71
D	0.67	0.77	0.66
平均	0.7	0.78	0.72

5.3 提示する評判情報の評価

ユーザに文単位で提示する評判情報について、実験により正解率を算出する。計算は、正解の評価視点に関して提示される評判情報のみ対象として行う。誤った評価視点に対しては正解の評判情報が提示されることはないため、計算の対象外とする。実験結果を表 4 に示す。

いずれのスコアリング指標においても、選択された評価視点について高い割合で正解の評判情報が提示されている。正解率が最も高かったのは、共起ボタンによりスコアリングした場合であった。出現頻度によるスコアリングにおいて提示した評判情報は、評価視点「部屋」に関する評判情報「また、部屋数に対して駐車場が不足していると思う」のように評判情報ではあるがその評価視点に対してではない場合、評価視点「レストラン」に関する評判情報「2 階のレストランにも行きたいです」のように評判ではなく願望を表す場合があったが、共起ボタンによるスコアリングではこれらを除去することができている。

6. 考 察

商品説明ページを用いた評価視点の抽出について、適合率が大幅に向上したことから、商品説明ページを用いた評価視点の絞りこみが有効に機能しているといえる。一方で、再現率は向上の余地があり今後の課題として残った。解決策として、抽出されなかった評価視点と抽出された評価視点との間にある関連性を統制することが考えられる。

商品説明ページに出現しない評価視点を提示する機能についても実装と評価をしたが、定量的な有効性を確認するには至らなかった。スコアリング手法の更なる改良とともに、利用者実験等も含めて評価を行う必要があるといえる。

本研究の評価実験では、閾値を 1 つの値に固定して行ったが、

閾値を下げることで適合率を維持したままで再現率を向上できる場合があった。閾値を変化させた評価を行うことで、利用者が閾値を柔軟に設定できるシステムの有効性を評価できると考える。また、人手で抽出した評価視点の正解データには、名詞以外の品詞を評価視点とするものも含まれていた。たとえば、「飲み放題」は、動詞「飲む」と名詞「放題」から構成される評価視点である。接続回数が少ないために、評価視点候補として抽出できていない複合名詞が評価視点となる場合もあり、抽出対象の品詞、複合名詞の抽出精度の向上は今後の課題である。

正解か否かの判断は、正解ラベルを付与された文字列との完全一致により行ったが、例えば「つくば駅に近くで便利」のような文には一貫して「駅」に対して正解ラベルを付与しているため「つくば駅」が評価視点として誤検出されていた。この点の改善も今後の課題である。

本研究で提案するシステムは、評判情報と商品説明ページがともに提供されているクチコミサイトであれば適用できることから、ホテルの評判情報以外の、例えば、パソコン、車、映画、レストランといった幅広い領域の対象物に適用できると期待される。また、評判情報を好評・不評に分類する技術と組み合わせることで、評価視点毎に好評・不評の割合を可視化し、対象物についてどこが良くどこが悪いのか俯瞰することができるなどの拡張も考えられる。

また、多くの人が注目する評価視点とそうでない評価視点がある。評価視点リスト中の評価視点を、言及回数により異なる文字サイズで提示することで、ユーザの評価視点選択を支援することができる。更に、先述した好評・不評の割合による文字色の決定と組み合わせることで、ユーザは評価視点の特徴を容易に把握することができる。

7. おわりに

Web上のクチコミサイトなどには商品・サービスに対する評判情報が大量に蓄積されている。評判情報は消費者が商品を購入するとき、企業が商品を改善するときの参考として有用である。しかし、対象物についての評判情報が膨大であれば、全てに目を通すには多大な負担がかかる。

本研究は、評判情報の閲覧支援を目的に、評判情報を評価視点別に分類し提示するシステムを提案した。多様な対象物に対応するために、評価視点は入力された評判情報から抽出した。評価視点の抽出に関する先行研究では、事前に作成した辞書を用いる手法や、評判情報記事中の語の出現状況を用いた手法が提案されている。しかし、前者の手法では辞書を作成するときにコストがかかり、後者の手法では実用レベルの精度、すなわち実用上十分なレベルの適合率が得られているとは言えないのが現状である。そこで、本システムは評価視点の抽出の精度向上を目指して商品説明ページを抽出の手がかりとして用いた。商品説明ページにはその商品の側面が網羅的に記述されている場合が多く、評価視点の抽出に利用することができる。抽出手順は、まず記事中の語の出現状況を用いて評価視点の候補をスコアリングし、そしてスコアが一定以上の候補のうち商品説明ページに出現する語を評価視点とした。

対象物に関する評価視点が複雑な構造を持つ場合を踏まえ、評価視点を商品説明ページと連動することで視覚的に構造化して提示し、ユーザの評価視点選択を支援した。また、抽出しなかった評価視点について、スコアに基づき整列し提示した。

提案システムについて、Web上で公開されているホテル予約サイトを対象として評価実験を行った。実験の結果、商品説明ページに出現する評価視点に限り提示することで、適合率が向上することを確認した。また、提示部に商品説明ページを用いることで、「レストラン」と「夕食」といったような関連の強い評価視点が近接して表示されるなど、視覚的に構造化され表示されることを確認した。

残された課題として、再現率を向上すること、ホテル以外の領域で提案手法の有効性を確認することがある。また、システムの利便性向上のため、評判情報を好評・不評に分類する機能、評価視点に関して好評・不評や言及回数を可視化する機能との組み合わせが考えられる。

謝 辞

本研究の一部は、「平成20年度 筑波大学 図書館情報メディア研究科 プロジェクト研究」の支援を受けて行われました。

文 献

- [1] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, pp. 79–86, 2002.
- [2] Michael Gamon, Anthony Aue, Simon Corston-Oliver, and Eric Ringger. Pulse: Mining customer opinions from free text. In *Proceedings of the 2005 Conference on Intelligent Data Analysis*, pp. 121–132, 2005.
- [3] 藤村滋, 豊田正史, 喜連川優. Webからの評判および評価表現抽出に関する一考察. 情報処理学会研究報告, 2004-DBS-134, Vol. 2004, No. 72, pp. 461–468, 2004.
- [4] 立石健二, 福島俊一, 小林のぞみ, 高橋哲朗, 藤田篤, 乾健太郎, 松本裕治. Web文書集合からの意見情報抽出と着眼点に基づく要約生成. 情報処理学会研究報告, 2004-FI-076, Vol. 2004, No. 93, pp. 1–8, 2004.
- [5] Mingqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177, 2004.
- [6] 飯田龍, 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. 意見抽出を目的とした機械学習による属性-評価値対同定. 情報処理学会研究報告, 2005-NL-165, Vol. 2005, No. 1, pp. 21–28, 2005.
- [7] 土田正明, 水口弘紀, 久寿居大. 対象-属性-評価の3項関係同定による評判情報抽出. 言語処理学会第13回年次大会発表論文集, pp. 412–415, 2007.
- [8] 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, 2004.
- [9] 阿辺川武, 奥村学. 形容詞を用いた対象-属性名詞対の収集および分析. 言語処理学会第12回年次大会発表論文集, pp. 73–76, 2006.
- [10] Christopher Scaffidi, Kevin Bierhoff, Eric Chang, Mikhael Felker, Herman Ng, and Chun Jin. Red opal: product-feature scoring from reviews. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, pp. 182–191, 2007.