

宿泊レビューを用いた宿泊施設サービス改善のための情報抽出

Information extraction for accommodations service improvement using user review

辻井 康一
Koichi, Tsujii

津田 和彦
Kazuhiko, Tsuda

筑波大学 システム情報工学研究科
Graduate School of Systems and Information Engineering, University of Tsukuba

Online hotel reservation is widely used for planning travel today. Users of reservation usually refer to the user reviews before making their reservations. Occasionally dissatisfied impressions and satisfied impressions are seen in the comments simultaneously. In this paper, we provide a method to analyze the relation of impressions for users using text mining. And we show the way of extracting useful information for hotel service improvement.

1. はじめに

インターネットの普及により、ホテル・旅館を予約する手段としてオンライン宿泊予約が主流となっている。オンラインで宿泊予約をする際には旅行代理店の店頭と違い、宿泊検討者が自らの判断での予約が必要となる。宿泊予約サイトでは宿泊利用者に対し宿泊後にアンケートを行い、利用者のプロフィール、数値評価項目、感想コメントの3つの項目からなる感想情報を宿泊レビューとして提供している。宿泊レビューは、実際に利用体験者の生の声であり、宿泊先を探そうとしている宿泊検討者にとって重要な選択判断材料となっている。

一方、宿泊サービスを提供する宿泊施設側もこのような状況を受け、宿泊レビューに書かれた評価内容に対し関心が高まり、自社の評判の把握やサービス向上改善の参考情報に利用するようになってきている[植竹 2011]。

宿泊施設がサービス向上を目指すためには、その施設が持つ具体的な問題点、改善点を洗い出すことが必要となる。これに対し、宿泊レビューの感想コメントには改善点の参考となる情報が利用者の目線で書かれている。しかし、感想コメントは人気ホテルになるほどその投稿数が膨大となり、個々の投稿を閲覧しサービス改善のための情報を得ることは非常に労力がかかる。そのため、効率的な情報収集の方法が必要である。

そこで本研究ではサービス改善のための参考となる情報を、宿泊利用者が投稿した宿泊レビューの感想コメントを用いて、テキストマイニングの手法により抽出する方法を検討する。

2. 宿泊施設の評価情報

宿泊施設の評価や改善点を見つけるための方法に関する研究として、大内らは、宿泊施設によるアンケート調査による評価手法の研究を行い、宿泊検討者への情報提供方法を検討している[大内 2008]。また辻井らは、宿泊レビューの感想コメントから、肯定評価と否定評価の特徴表現を抽出し、数値評価項目と対応させることでエリアの特徴、宿泊施設の特徴を表現する方法を示した。また評価表現は宿泊利用者の関心の高い数値評価項目に集中することや、人気宿泊施設において肯定的な評価表現の一部に否定的な評価表現も書かれていることを示した[辻井 2012]。

感想コメントの文中に出現する否定的な評価表現は、宿泊利

連絡先: 辻井 康一, 筑波大学 システム情報工学研究科
〒112-0012 東京都文京区大塚3丁目29-1,
e-mail : tsujii@gmail.com

用者が体験した内容に基づくものであるが、否定評価や一部満足できないことがあっても、全体としては満足といった肯定評価と否定評価が同時に記され、評価極性が遷移する表現が見られる。この表現から、満足を感じつつも同時に不満足である、重点的に改善すべき項目を明らかにすることができる可能性がある。これらはサービスを改善するために必要な要素を提供する有益な情報といえる。これらの評価が文中で遷移する評価表現から肯定、否定評価部分を抽出し、宿泊施設のサービス改善につながる特徴項目を発見することを目指している。

3. サービス改善のための評価表現

サービスの改善を目指すための情報を抽出するため、利用者が体験し感じた感想コメントから改善対象項目を抽出しなければならない。そのため宿泊利用者が投稿する感想コメントのテキスト情報を利用した評価表現の抽出方法を検討する。

3.1 評価表現の抽出

感想コメントを文単位に見た場合、下記のように文の前後において評価表現が異なって記載されているものがある。

- ①部屋は狭かったけど、料金が安かったので満足です。
- ②料理はおいしかったけれど、量が多すぎです。

①の場合は前半が否定評価、後半が肯定評価となり、②の場合は、前半が肯定評価、後半が否定評価となっている。尾括型傾向の強い日本語では、後半の評価表現の高く評価する。前半の評価表現は後半の評価ほど重要度は高くはないが、比較する対象となった特徴的な項目である可能性を示している。また、下記のように評価表現の対象と評価の組み合わせが異なる場合も見られる。

- ③部屋は狭いのだが、部屋の雰囲気は最高 (対象同一)
- ④風呂は残念だけでも、部屋が広いから満足 (対象違い)

③はどちらも部屋を話題とした評価表現であり部屋に対する「広さ」と「雰囲気」という評価属性を比較し評価している。④は前半部分が風呂に対する評価であり、後半部分は部屋に対する評価となる。この場合は、宿泊利用者は宿泊体験の総合的な判断をもとに評価を行っていると考えられる。また、評価対象となる事柄は多岐にわたるため、辞書構築による取りまとめが必要になる。

3.2 辞書の登録

感想コメントは宿泊利用者が自由に記述する情報のため、表現の揺れが大きい。そのため類義語辞書を構築し、表現を揃える処理を行う。次に評価表現の肯定、否定極性を判定するために評価表現辞書を構築する。その際、本研究では小林らの評判分析に用いる肯定・否定表現の対象、属性、評価の3つ組構成を活用した[小林 2005]。数値評価との連携を考慮し<対象>は数値評価項目とし、<属性>を“大浴場”や“ベッド”など具体的な評価物、<評価>を評価物とする構造とした。

これらの方法をもとに、感想コメントの係り受け解析を行い、肯定から否定、否定から肯定に評価が遷移するものを抽出する。感想コメントから利用者が実際に体験した事実を評価した肯定表現、否定表現を抽出する事で、宿泊者が感じたサービスの現状を知ることが出来る。これらを集計し、その特徴を示す事でサービス提供者である宿泊施設に対し、改善項目を検討するための情報提供が出来ると考えられる。

4. 評価表現の抽出

前章で検討した手法を用いて、評価が遷移する評価表現を抽出する。データは宿泊予約サイトのじゃらん.net¹の宿泊レビュー情報を利用した。ビジネス用途、レジャー用途のホテル、旅館の感想コメントを形態素解析し、文単位で、「けど、だけど、だが」といった文中に逆接の接続助詞のあるものを抽出する。それらの係り受け解析を行い、前半部分と後半部分において評価表現が遷移する文のみを抽出し、分析対象とした。

4.1 改善すべき対象項目の抽出

本節では評価対象となる数値評価項目と評価の肯定・否定極性の関連を検討する。文中の評価が否定表現から肯定表現に遷移したものを集計し、その頻度を表 1 に、反対に肯定表現から否定表現に遷移したものを表 2 に示す。

表 1. 否定表現から肯定表現への遷移

否定→肯定	サービスP	施設P	清潔感P	部屋P	風呂P	立地P	料金P	料理P	総計
サービスN	3					1			4
施設N		11	6	1			2		21
清潔感N			1						1
部屋N		2	8	26	3		2	1	42
風呂N		1	2	1	12	1	1	1	19
立地N		1	1			4			8
料金N		1	2		1		4		8
料理N		1	1			1		12	16
総計	8	19	14	30	17	6	10	15	119

表 2. 肯定表現から否定表現への遷移

肯定→否定	サービスN	施設N	清潔感N	部屋N	風呂N	立地N	料金N	料理N	総計
サービスP	1			2					3
施設P		2						1	3
清潔感P									
部屋P		1		6			1		8
風呂P					9				9
立地P		1		1		1			3
料金P		1	1	2			1		4
料理P		1	1					17	19
総計	4	4		11	9	1	2	18	49

表側は前半の評価を示しており、N は否定評価を、P は肯定評価を示している。表頭は後半の評価を示している。両表から宿泊利用者は、否定表現から肯定表現への評価の遷移のほうが肯定評価から否定評価への遷移と比べて多いことがわかる。また同一の評価対象項目に対し肯定評価、否定評価を行っている傾向が強い。次に、表 1 を見ると、部屋に関する否定表現から肯定表現への遷移が多くみられた。部屋に対する否定的な評価を持つが、後半の評価により肯定的な判断に移ることがわかる。この場合、宿泊利用者は感想コメントを記載する際、前半

部分の評価項目に不満を持ちつつも、後半部分の評価項目に満足している状態と想定される。前半の評価部分を改善することで、利用者の満足度をより高められると考えられる。

表 2 は肯定表現から否定表現への遷移となるが、表 1 と比べ頻度が少なく、部屋、風呂、料理についての評価の遷移が多くみられた。後半部分の評価内容が利用者の満足度を下げている要素である可能性がある。

4.2 宿泊利用用途の違いによる特徴

表 1 の結果をビジネス利用とレジャー利用の施設に分類し、否定評価から肯定評価へ遷移する表現の傾向を表 3 に示す。

表 3. 宿泊施設タイプによる否定表現から肯定表現への遷移

否定→肯定	サービスP	施設P	清潔感P	部屋P	風呂P	立地P	料金P	料理P	総計
ビジネス	2	7	7	21	7	2	5	4	55
サービスN									1
施設N	1	2	1						4
清潔感N					1				1
部屋N		1	6	18	2		2		29
風呂N		1	1	1	4				6
立地N				1		1			2
料金N	1	2		1			2		6
料理N		1					1	4	6
レジャー	6	12	7	9	10	4	5	11	64
サービスN	3								3
施設N		9	5	1			2		17
清潔感N									0
部屋N		1	2	8	1			1	13
風呂N	1	1			8	1	1	1	13
立地N		1	1			3		1	6
料金N							2		2
料理N	1				1			8	10
総計	8	19	14	30	17	6	10	15	119

ビジネス利用の施設では、部屋に関する内容が一番多くみられた。具体的には部屋の広さに不満があるが、広さ以外の属性要素への満足度が高い事を評価する表現が多くみられた。レジャー利用の施設では評価対象が分散し、施設、部屋、風呂、料理の評価表現において評価の遷移が見られ、ビジネス利用の際と同様に同一対象の別属性への評価がされる傾向がある事がわかった。これらの結果より、宿泊施設がサービスを改善する場合には、宿泊施設の利用用途に合わせて、同一評価対象の異なる属性の評価向上および改善に注力することで、宿泊利用者の評価を効率的に高めることができる可能性を示している。

5. おわりに

本研究では、宿泊レビューの感想コメントを用いた評価対象への肯定表現と否定表現の遷移に着目した評価表現の抽出方法を提示した。宿泊施設の評価表現は同一評価対象において評価が遷移すること傾向が見られた。評価の遷移の中から不満の原因となる評価表現を抽出することで、宿泊施設のサービス改善につながる項目を明らかにする可能性を示した。

今後の取り組みとして、評価情報収集の効率化の検討を行うとともに、サービス改善につながる可能性のある対象項目の宿泊施設への情報提示方法の検討を行う。

参考文献

[植竹 2011]植竹 朋文,青木 章通, “繁閑格差の大きい地区のホテルにおけるネット直販に関する意識調査”, 専修大学情報科学研究所所報 No.76 pp.1-26 (2011).
 [大内 2008]大内 東,大柳 幸彦:日本のホテル評価システムに関する研究,開発こうほう, (539), pp.27-31 (2008).
 [辻井 2012]辻井 康一,津田 和彦, 2012. テキストマイニングを用いた宿泊レビューからの注目情報抽出方法. 情報処理学会デジタルプラクティス, 3(4),pp.289-296.
 [小林 2005]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 2005

¹じゃらん net <http://www.jalan.net>