

ユーザーレビューを用いた宿泊施設の消費者選好分析に関する研究

後藤 正幸 研究室

0532027 井上 淳

1. 研究背景・目的

近年、インターネットの普及によるサービスの変化は、宿泊業界にも大きな影響を与えている。その一つが、宿泊予約サイトの出現である。これまで、宿泊施設に直接電話、または旅行代理店に出向くことで宿泊予約をすることが一般的であった。しかし、宿泊予約サイトの普及によって、インターネット上で宿泊施設の空室状況を調べ、予約することが可能となった。また、宿泊予約サイトには、宿泊施設の利用者がコメントによる評価を残すことができ、多くの消費者からの情報が寄せられることで、様々な視点からの評価が集積されている。先行研究[1]では宿泊予約サイトの利点を活かし、投稿されたユーザーレビューから、「オケーション」、「部屋」、「食事」、「お風呂」、「接客・サービス」に分類・整理を行い、宿泊施設の評価指標である48項目を作成し、これらの項目に対して、270人分のユーザーレビューから、各々の宿泊施設の特徴を分析している。その結果、利用者への満足度が高い宿泊施設の要因が明らかとなり、ユーザーレビューから有用な知見を得られることがわかっている。しかし、消費者の性別や年齢などの消費者属性別に期待される宿泊施設の選考要因、宿泊施設に対する総合満足度評価による差異は明らかとなっていない。また、全て手作業による分析のため、分析データ数の大幅な増加が難しいなどの課題があった。

そこで本研究では、人手ではコスト面から効率的な活用が困難であった大量のユーザーレビューの解析に対して、テキストマイニングの手法を適用する。この手法により、大量のテキストデータから短時間で有用な情報を得ることが可能となる。具体的には、ユーザーレビューに含まれる単語の出現頻度を測定し、統計的な指標値を算出することで、消費者の属性別、総合満足度評価別の特徴的な宿泊施設の個別評価項目を明らかにすることを目的とする。これによって、宿泊施設に対する消費者の選好要因を詳細に把握でき、個々の宿泊施設の強みを活かした戦略を立案することが可能となる。

2. 研究手順

①宿泊予約サイト上から抽出したユーザーレビューに対し、形態素解析を行い、単語の出現頻度を測定する。②ユーザーレビューよりレビュー属性、単語の出現頻度上位から宿泊施設の個別評価項目を抽出する。③各レビュー属性における特徴的な宿泊施設の個別評価項目を、補完類似度[2]を用いて算出・比較する。④抽出された特徴的な宿泊施設の個別評価項目の係り受け分析を用いて、詳細に把握する。

3. 研究内容

3-1. 分析対象

本研究では、宿泊予約サイト「じゃらんnet」(<http://www.jalan.net/>)内のユーザーレビューを用いる。分析対象は、2008年7月26日～8月17日の期間で「じゃらんnet」内にある、「ユーザークチコミ宿ランキング」の「泊まって良かった(総合)宿ランキング」、「お風呂の良かった宿ランキング」、「夕食の良かった宿ランキング」、「お部屋の良かった宿ランキング」の関東・甲信越エリアの「旅館」における、各ランキング上位10件、計2,246件のユーザーレビューとする。

3-2. テキストマイニングによるユーザーレビューの解析

ユーザーレビューの自由記述文章に対して、形態素解析と構文解析を行った。形態素解析から単語頻度分析を行い、宿泊施設の個別評価項目の抽出を行う。次に係り受け分析を行い、係り受け関係の推定を実行した。なお、本研究では株式会社数理システムのテキストマイニングツールである「Text Mining Studio」を利用した。

3-3. 宿泊施設の個別評価項目・各属性の定義

3-3-1. 宿泊施設の個別評価項目の定義

「単語頻度分析」の結果より、上位30件を対象として、宿泊施設の全体的な印象である「宿」、「清潔感」、各宿泊施設の特徴が最も現れると考えられる「食事」、「部屋」、「お風呂」、人と人とのコミュニケーションによる印象である「スタッフ」、「宿主」、「気遣い」、「サービス」、観光地・名所を重視していることを示す「地名」、「景色」、宿泊施設特有な項目である「貸し切り」、の宿泊施設と関係があると思われる名詞、12個を選定し、「宿泊施設の個別評価項目」として定義した。

3-3-2. 消費者属性の定義

ユーザーレビューの基本情報からは、「性別」、「年代」、宿泊した際の価格帯を示す「価格帯」、どの時期に行ったかを示す「季節」の4個を抽出した。本文からは、書いてあったものには○、書いていなかったものについては×として、何のために行ったのかを意味する「イベント」、誰で行ったのかを示す「同伴者」、複数回訪れたことを書き込んである「リピーター」、クチコミを参考にして宿を決定した「クチコミ」の4個を抽出し、計8個を、「消費者属性」と定義した。

3-3-3. 総合満足度評価の定義

ユーザーレビューに宿泊後の印象として、宿泊施設に対する満足度に影響している評価項目の抽出が可能であると考えられるため、「また行きたい」、「オススメ」を抽出し、計2個を「総合満足度評価」として定義した。

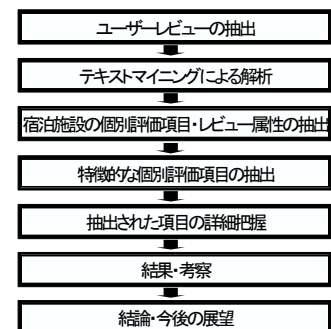


図1 研究手順

3-3-4. レビュー属性の定義

「消費者属性」、「総合満足度評価」の2つをまとめたものを「レビュー属性」として定義した。

3-4. 統計的指標値を用いた特徴的な個別評価項目の抽出

宿泊施設に対する消費者の選好要因を明らかにするため、「消費者属性」、「総合満足度評価」ごとの特徴的な「宿泊施設の個別評価項目」を抽出する。しかし、各レビュー属性の単語の出現頻度が異なるため、単純に頻度をそのまま比較をすることができない。そこで本研究では、特徴的な「宿泊施設の個別評価項目」を抽出するために、補完類似度[2]を算出し、補完類似度が高いものを特徴的な宿泊施設の個別評価項目とし、比較を行う。補完類似度とは、単語の出現頻度を基に、対象となる単語の全体における出現頻度とその属性における出現確率を考慮して、特徴指標値を算出する方法である。

今回は複数のレビュー属性間での比較も行うため、各レビュー属性における各宿泊施設の評価項目の出現割合を算出してから、補完類似度を求める。これにより、算出した補完類似度がレビュー属性内だけではなく、レビュー属性間や全体に対しての意味も持たせることができる。補完類似度の計算のために必要な変数の定義を以下に示す。

P_{r_j, w_i} = 属性 r_j における項目 w_i の出現割合

上の変数を10属性×12項目について求め、表2を完成することで、(1)式に示す補完類似度が計算できる。

3-5. 特徴的な個別評価項目の詳細把握

抽出した特徴的な宿泊施設の個別評価項目は、それだけでは何について書かれているのかがわからない。例えば、「清潔感」という項目が抽出されても、何についての「清潔感」なのかは把握できていないため、「お風呂についての清潔感」といったように文脈で判断する必要がある。そこで「係り受け分析」を用いて、「清潔感-お風呂」のような係り受け表現を抽出する。その結果から上位に出現した表現を用いて、何について書かれているのかを把握する。また、「係り受け分析」でも係り元と係り先が逆になっている場合、「清潔感-お風呂」、「お風呂-清潔感」といったものは同じ意味として捉え、まとめて1つとして考えるため、グループ化を行い、分析する。

4. 結果

紙面の都合上、総合満足度評価の「また行きたい」における各個別評価項目の補完類似度の算出結果を示す(図2)。その結果、宿泊施設の主要な評価項目である、「食事」、「お風呂」といった項目が低く、「気遣い」、「宿主」、「サービス」、「清潔感」、「地名」などの項目が高かった。次に、抽出された評価項目である、「気遣い」、「宿主」、「サービス」、「清潔感」、「地名」に対して係り受け分析を行った。分析結果の上位となった、「気遣い」では、利用者に対する細かい気遣いと考えられる「気遣い-細かい」、隅々まで行き届いた気遣いと考えられる「気遣い-行く」などが得られた。「宿主」では、「宿主-気遣い」、「宿主-人柄」といった人と人とのコミュニケーションにつながる表現が上位であった。「サービス」は、「サービス-丁寧」、「サービス-細かい」といったサービスについての印象が上位であった。「清潔感」では、宿泊施設の全体的な印象である「清潔感-宿」、次に「清潔感-部屋」が上位であったことから、部屋の清潔感の印象が重要な要因であることが考えられる。「地名」では、観光地を訪れたことを示す「地名-行く」、観光地で見たを意味する「地名-見る」という結果が得られた。

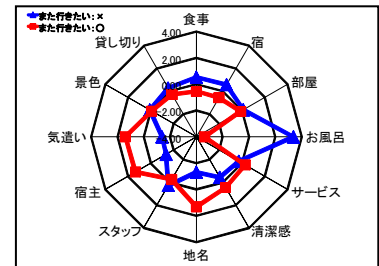


図2 補完類似度結果: また行きたい

5. 考察

ユーザーレビューに「また行きたい」と書き込んだ消費者は、次も利用したいという気持ちが高いと考えられ、リピーターにつながる可能性が高いことが考えられる。つまり、リピーターにつながる要因の把握が可能であり、補完類似度が高い「気遣い」、「宿主」、「サービス」のような人と人とのコミュニケーションによる評価項目は注目すべき点である。すなわち、これらの項目に力を入れることで消費者に好印象を与え、満足度を高めることが可能と考えられる。一方、「お風呂」の評価項目が極端に低いが、これは同じように結果が低かった「部屋」、「宿」といった項目には、「清潔感」での係り受け表現の上位に出現しており、間接的な関係があると考えられる。しかし、「お風呂」の評価項目は間接的な関係がなく、宿泊施設に対する1つの独立した評価基準となっている可能性が考えられる。

6. 結論

本研究では、宿泊予約サイト上のユーザーレビューから、各レビュー属性の各宿泊施設の個別評価項目ごとの出現割合を用いて補完類似度の算出・比較を行い、特徴的な個別評価項目を明らかにした。その結果、先行研究では解明されていなかった消費者の属性、総合満足度評価別による特徴的な宿泊施設の個別評価項目の差異を明らかにすることができた。これにより、宿泊施設が個々の強みを活かした戦略の立案が可能であると考えられる。

7. 今後の展望

本研究によって、宿泊施設が戦略を構築する際に必要な、消費者属性、総合満足度評価別の特徴的な宿泊施設の個別評価項目が明らかとなった。しかし、特徴的な個別評価項目の抽出はできたが、個別評価項目についての詳細な把握がまだ十分ではない。そのため、例として「お風呂」の何について力を入れるべきなのか、個室風呂あるいは露天風呂なのかというアプローチ方法までは示せていない。よって、個別評価項目の類義語について分析を行うことでより詳細な戦略の立案が可能であると考えられる。

参考文献

- [1] 田邊 亘:「宿泊施設の戦略構築を支援するユーザーレビュー分析に関する一考察」, 武蔵工業大学環境情報学部, 情報メディアセンタージャーナル, Vol.9, pp.91-101, (2008)
- [2] 株式会社数理システム, 「Text Mining Studio 技術資料」