日本ソーシャルデータサイエンス学会
2026 年春季研究発表会

# 宿泊施設レビューデータを用いた顧客満足度の構造分析

園　亮太朗†1，大場　春佳 1 ，生田目　崇 1

**キーワード**：レビューデータ，TOBIAS，BERT モデル，PLS-PM

## 1.　目的

近年，地方観光振興において，観光客の多様化や競争の激化を背景に，観光体験や宿泊施設におけるサービスの向上が求められている．特に，利用者によって投稿されるレビューデータは，観光客の意思決定や施設改善・経営戦略を支える重要な情報源となっている．実際に，宿泊施設選定の際，旅行情報サイトや口コミサイトを参考にする利用者は多い[1]．そこで，本研究では宿泊施設のレビューデータを用いて，観光地における宿泊施設の満足度の構造や要因を明らかにすることを目的とする．

## 2.　使用するデータ

本研究では，国立情報学研究所の IDR データセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」(https://rit.rakuten.com/data_release/)の楽天トラベルレビューデータ[2]を使用した．

・対象期間：2014 年 1 月 1 日～2019 年 12 月 19 日
・対象地域：兵庫県
・対象施設：74
・データ件数：35,880 件

## 3.　分析方法

本研究では，TOpic modeling Based Index Assessment through Sentiment (以下 TOBIAS)[3]を用いて分析を行った．TOBIAS はレビューに含まれるトピックや感情が，満足度や評価にどのように影響しているかを，感情評価，トピック抽出，PLS-PM というステップで分析を行う．以下の図 1 に分析の流れを示す．

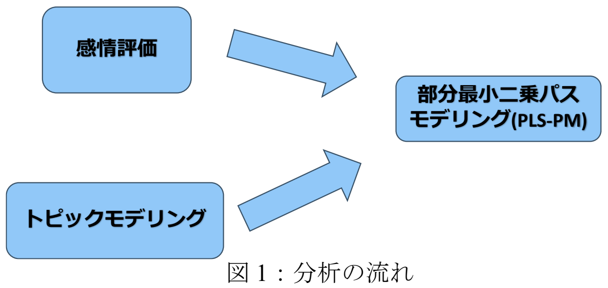


図 1：分析の流れ

本研究では感情評価において，各レビューに含まれる感情的特徴を定量化するために，感情極性の推定に BERT を，感情の種類に関するスコア推定に DeBERTa[4]を用いた．また，各レビューが属するトピック情報を PLS-PM 分析における観測変数として用いるために，トピック抽出において，BERTopic[5]を用いた．そして，得られた感情指標や，トピック情報が満足度に与える影響を推定するために，PLS-PM を用いた．

## 4.　結果・考察

### 4.1 兵庫県全体の結果

ここでは，兵庫県の宿泊施設のレビューを対象に PLS-PM を行なった結果について述べる．まず，感情評価の結果，感情極性の推定において，約 8 割がポジティブと判定された．また，感情の種類の推定では，「喜び」「期待」「信頼」「驚き」「悲しみ」「怒り」「恐れ」「嫌悪」の 8 つの感情を判定した．その結果，「喜び」が最も多く判定された．

次に，トピック抽出では，レビューを 10 のトピックに分類した．以下の表 1 は各トピックの内容を示したものである．

表 1：トピック抽出の結果

| トピック番号 | 内容 |
|---|---|
| 0 | ホテルの設備・サービスについて |
| 1 | ホテルの設備・サービスについて，記念日や家族旅行での利用が多め，サービスについて高評価多め |
| 2 | ホテルの設備・サービスについて，フロント対応や設備について不満も目立つ |
| 3 | ホテルの設備・サービスについて，風呂，温泉について述べている，不満も目立つ |
| 4 | 神戸観光または出張などの利用 |
| 5 | 淡路島観光や有馬温泉観光のレビューが多い |
| 6 | 神戸の中でも特に三宮にフォーカスしている |
| 7 | 姫路駅，姫路城周辺施設のレビュー |
| 8 | 甲子園やキッザニア利用者多め |
| 9 | 城崎温泉の外湯めぐりや料理への満足度 |

表 1 より，トピック 0,1,2,3 はホテルの立地や朝食などのサービスについての話題が多いことが分かる．また，トピック 4,6 では神戸市に関する話題が多いことが分かる．特にトピック 6 は神戸市の中でも三宮に関する話題が多い．一方で，トピック 5,7,8,9 は，淡路島や姫路など神戸市以外の地域の話題が出現していることが分かる．

そして，感情評価による感情スコアの推定値とトピック抽出によるトピック情報，総合評価値や施設のサービスや設備に対する評価値を観測変数として用いて，PLS-PM を行った．パス図は以下の図 2 の通りである．

†1 中央大学

ⓒ2026 Japan Social Data Science Society

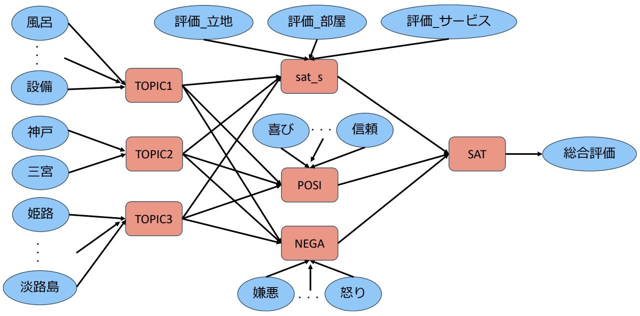


図 2：PLS-PM のパス図

ここで，青色で示したものは観測変数であり，赤色で示したものは潜在変数である．兵庫県全体のレビューを対象に PLS-PM を行った結果のパス係数の一部を以下の表 2 に示す．

表 2：PLS-PM によるパス係数の結果

| 内側モデル | | 外側モデル | |
|---|---|---|---|
| パス | パス係数 | パス | パス係数 |
| TOPIC1→sat_s | 0.2736 | トピック2→TOPIC1 | 0.6843 |
| TOPIC1→POSI | 0.3501 | トピック3→TOPIC1 | 0.7348 |
| TOPIC1→NEGA | -0.3778 | トピック4→TOPIC2 | 0.8818 |
| TOPIC2→sat_s | 0.0382 | トピック6→TOPIC2 | 0.5440 |
| TOPIC2→POSI | 0.0760 | トピック5→TOPIC3 | -0.6126 |
| TOPIC2→NEGA | -0.0439 | トピック7→TOPIC3 | 0.7332 |
| TOPIC3→sat_s | 0.0510 | 評価_立地→sat_s | 0.1447 |
| TOPIC3→POSI | 0.0860 | 評価_部屋→sat_s | 0.4934 |
| TOPIC3→NEGA | -0.0936 | 評価_サービス→sat_s | 0.5509 |
| sat_s→SAT | 0.6482 | 喜び→POSI | 0.6987 |
| POSI→SAT | 0.1024 | 信頼→POSI | 0.3234 |
| NEGA→SAT | -0.1899 | 怒り→NEGA | 0.1340 |
| | | 嫌悪→NEGA | 0.3758 |

ここで，内側モデルは潜在変数間の関係，外側モデルは観測変数と潜在変数間の関係を表す．表 2 より，外側モデルでは，「怒り」よりも「嫌悪」の方が，数値が高いことが分かる．また，「悲しみ」も「怒り」より数値が高い結果であった．このことから，ネガティブ感情は怒りよりも落胆や不快感によって構成されていると考えられる．また，内側モデルでは，顧客満足度(SAT)に大きく影響するのは，サービスや部屋への満足度(sat_s)であることが分かる．一方で，感情要因の影響については，ポジティブ感情よりもネガティブ感情の方が，パス係数の絶対値が大きいことが分かる．このことから，満足度の維持には，感動の提供以上に，ネガティブ感情の回避が重要であると考えられる．

### 4.2 地域ごとの結果

また本研究では，兵庫県を「淡路」「播磨」「但馬」「摂津」の 4 地域[6]に分け，再度 PLS-PM を行った．ここでは 4 地域の一部の結果を以下に示す．

表 3：地域ごとの PLS-PM によるパス係数の結果

| パス | パス係数_淡路 | パス係数_播磨 | パス係数_但馬 | パス係数_摂津 |
|---|---|---|---|---|
| sat_s→SAT | 0.6345 | 0.6422 | 0.6273 | 0.6550 |
| POSI→SAT | 0.1184 | 0.1118 | 0.1331 | 0.0946 |
| NEGA→SAT | -0.1729 | -0.1951 | -0.1554 | -0.1906 |

表 3 より，すべての地域において，顧客満足度(SAT)に最も影響するのは，サービスや部屋への満足度(sat_s)であることが分かる．また，感情が顧客満足度に与える影響を比べてみると，都市部に位置する摂津エリアではポジティブ感情の影響に対し，ネガティブ感情の影響は約 2 倍の大きさを示している．このことから，摂津エリアは，顧客がサービス品質に対して高いベースラインを求めており，サービス品質の低下が満足度の低下に影響しやすいと考えられる．そのため，顧客に感動を与えることよりも，サービス品質の低下を防ぎ不満を抑制することで，満足度を維持できると考えられる．

一方で，温泉地や豊かな自然に囲まれた但馬エリアでは，ポジティブ感情のパス係数が 4 地域の中で最も高く，ネガティブ感情の影響も約 1.2 倍であり，摂津エリアと比べると低くなっていることが分かる．つまり，但馬エリアにおいてネガティブ感情の影響が相対的に低下していることを示している．このことから，不満の抑制にとどまらず，非日常体験による感動の創出が満足度を高める要因になると考えられる．

## 5. まとめ・今後の課題

本研究では，兵庫県のレビューデータを対象に，TOBIAS を用いて，満足度の構造や要因を明らかにした．その結果，地域を問わずサービスや部屋への満足度が顧客満足度に最も影響する要因であることが分かった．また，兵庫県を 4 地域に分けて分析を行ったことで，地域特性によって感情要因の影響が異なることが分かった．特に観光地においては，不満の抑制だけでなく感動体験の提供が満足度向上において必要不可欠となる．

また，本研究では，対象がレビュー本文だけであったため，利用目的や同伴者の有無による差異を考慮できていない．そのため，今後の課題としては，利用目的や顧客属性を考慮した分析を行うことが挙げられる．

## 謝辞

本研究を進めるにあたり，データ提供をいただいた国立情報学研究所様及び，楽天グループ株式会社様にこの場を借りて厚く御礼申し上げます．

## 参考文献

[1] VEGAN’S LIFE , 「ホテル選びに関する調査 2024　byVEGAN’S LIFE」, https://vegans-life.jp/article/52_enpkqr5y　(閲覧日 2025-01-20)
[2] 国立情報学研究所，「情報学研究データリポジトリ　楽天データセット」, https://www.nii.ac.jp/dsc/idr/rakuten/ (閲覧日 2026-01-09)
[3] Marco Ortu, Luca Frigau & Giulia Contu, “Topic based quality indexes assessment through sentiment”, Vol. 39, pp.289–311, (2024)
[4] Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen, “DeBERTa: Decoding-enhanced BERT with Disentangled Attention”, arXiv preprint arXiv:2006.03654 (2020)
[5] Maarten Grootendorst, “BERTopic: Neural topic modeling with a class-based TF-IDF procedure”,  arXiv preprint arXiv:2203.05794 (2022)
[6] 兵庫県観光サイト HYOGO!ナビ，「五国エリア情報」, https://www.hyogo-tourism.jp/area/ (閲覧日 2025-12-11)