

## レビューテキストからの観点別評価点の推定

## Aspect-Based Rating Estimation from Review Texts

西垣 佑哉<sup>†</sup> 吉田 稔<sup>†</sup> 松本 和幸<sup>†</sup>  
Yuya Nishigaki Minoru Yoshida Kazuyuki Matsumoto

## 1. はじめに

近年、商品やサービスの評価情報がインターネット上で爆発的に増加しており、それについてのレビュー情報を見たいユーザーは大量のレビューを読む必要がある。しかし、これらのテキストは膨大であり全ての詳細を把握することは難しく、情報過多の中での要約と理解は難しいと考えられる。そこで点数での評価を利用することで、テキストによる評価に比べて素早くその商品やサービスについて理解できると考えられる。

また、レビューテキストには商品やサービスに対して様々な側面への意見や感情が豊富に含まれており、これらを的確に抽出し理解することが求められる。総合評価点だけでは、具体的な良い点や改善点などが不透明である。特定の側面に対する評価を主観的なテキスト情報から観点別に分類し、点数という形で可視化することで、ユーザが知りたい側面における情報が素早く理解できると考えられる。

## 2. 研究目的

本研究の目的はレビューテキストを観点ごとに分類して評価点を推定することである。楽天トラベル<sup>[1]</sup>におけるレビューテキストを6つの観点(立地、部屋、食事、風呂、サービス、設備・アメニティ)に分類して、観点ごとに評価点を推定する。SNSでの宿泊施設レビューでは、評価点という指標を持って評価されていることが少なく、大半はテキストによる評価のみである。それに対して、楽天トラベルのレビューでは、テキストでの評価と共に評価点という数字での評価も付けられている。これより、本研究での評価点を推定するモデルの学習には楽天トラベルのレビューデータを使用する。本研究の手法により高い精度で評価点を推定するモデルが実現すると、テキストのみで評価されているSNSの宿泊施設レビューテキストに、観点別の評価点という素早くかつ的確に理解することが容易である指標を付与することができると考える。

また、研究目的を達成することで、ユーザーが宿泊施設を利用する際に重要視する観点の評価や、ユーザーからの評価が低く事業者が改善すべき観点など、ユーザーと事業者の両方にとって有益な情報を提供することができると考える。

## 3. 関連研究

張らの研究<sup>[2]</sup>では、与えられたレビューテキストから評価視点毎にユーザの評価点を推定している。「部屋」に関する類義語などキーワードのリストを使って、レビュー全体から各項目に関連するコメント文を抽出し、そのコメント文をL2正則化ロジスティック回帰で分析し、評価点を予測している。本研究では、文脈を考慮した双方向の学習が可能であるBERTモデルを使用することで、文脈依存性や単語の関係性を豊富に捉えることができるため、高い精

度での評価点推定ができると考える。また、投稿レビューごとに全ての観点に関する評価点を推定するのではなく、レビューテキスト中に特定の観点に関する記述が見られる場合のみその評価点を推定する。

## 4. 研究手法

## 4.1.1 使用するデータ

本研究では、楽天データセット<sup>[3]</sup>の楽天トラベルレビューデータを使用している。このデータには、宿泊施設利用概要等が記載されているデータ(以下データ1とする)とレビュー評価が記載されているデータ(以下データ2とする)が存在する。各データに含まれている内容は下記の通りである。

- ・データ1: 施設番号, 投稿日時, ユーザー投稿本文, 投稿番号, 分類, プランID, プランタイトル, 部屋種類, 部屋名前, 施設回答本文
- ・データ2: 投稿番号, 投稿日時, ニックネーム, 目的, 同伴者, 参考になった数, 参考にならなかった数, 評価1(立地), 評価2(部屋), 評価3(食事), 評価4(風呂), 評価5(サービス), 評価6(設備・アメニティ)

## 4.1.2 レビュー文の分割

各観点に分類するために、1つのユーザー投稿本文を複数の文に、句点と逆説表現で分割する。逆説表現は文頭以外で使われると考えられる[ですが、たが、でしたが、だが、けど、だけど、ますが、けれど、対して、ものの、にも関わらず、にもかかわらず、とはいえ、いえども]を対象とする。逆説表現の置き換えの例を表1に、句点と逆説表現での分割例を表2に示す。

表1 逆説表現の置き換えの例

置き換え前のテキスト	置き換え後のテキスト
ですが	です
でしたが	でした
よかったですはいえ	よかったです

表2 レビュー文の分割の例

分割前のテキスト	分割後のテキスト
空調音が気になりましたが、	空調音が気になりました
食事は美味しくで大満足です。次回も利用したいです。	食事は美味しくで大満足です
	次回も利用したいです

このように分割することで、同一の文内に対照的な評価表現(Aは良かったがBは悪かった等)を存在しないようにしている。これは、評価点推定の際に誤った推定を行わないようにするためである。

<sup>†</sup>徳島大学大学院 Tokushima University

#### 4.1.3 観点分類用の観点辞書を作成

154,799 件のレビューデータより、MeCab [4] を用いてテキストから名詞のみを抽出して、頻出単語上位 3,000 件に対して観点ラベルを付与する。頻出単語への観点ラベル付与の例を表 3 に示す。

表 3 頻出名詞への観点ラベル付与の例

出現頻度順位	単語	出現数	観点ラベル
1	部屋	102,427	部屋
2	ホテル	61,410	観点なし
3	朝食	43,021	食事
:	:	:	:
2,998	外側	87	観点なし
2,999	横断	86	観点なし
3,000	叔母	86	観点なし

「ホテル」、「叔母」など、どの観点にも相応しくないと考えられる単語には観点ラベルを付与しない。

#### 4.1.4 観点の分類

観点辞書に存在する単語が、分割済みのレビュー文中に含まれていれば対応した観点ラベルを付与する。その例を表 4 に示す。

表 4 観点辞書を参照した観点ラベルの付与

分割済みレビュー文	辞書の単語	観点ラベル
設備や見た目を気にし…	設備	設備・アメニティ
部屋は新しいだけあつ…	部屋, 清掃	部屋・サービス
屋内浴室はもう少しア…	浴室	風呂

## 5. 実験

評価点推定の実験として 2 つの実験を行う。実験は東北大学の訓練済み日本語 BERT<sup>5)</sup> をファインチューニングして行い、モデルは各観点における精度・F1 スコア等にて評価する。

#### 5.1.1 評価推定実験 (5 段階)

楽天トラベルのレビュー評価尺度に沿って、ユーザーが付けている 1-5 の 5 段階の評価点を推定する。モデルごとの精度、macroF1 スコア、モデル平均適合率、モデル平均再現率を表 5 に示す。

表 5 5 段階評価推定実験の結果

モデル	精度	F1 スコア	適合率	再現率
立地	0.58	0.54	0.52	0.55
部屋	0.61	0.55	0.57	0.54
食事	0.60	0.62	0.60	0.62
風呂	0.48	0.45	0.46	0.42
サービス	0.54	0.50	0.52	0.49
設備・アメニティ	0.63	0.57	0.62	0.56
平均値	0.57	0.54	0.55	0.53

#### 5.1.2 評価推定実験 (3 段階)

ユーザーが付けている 5 段階の評価点より、評価点 1・2 を low, 評価点 3 を middle, 評価点 4・5 を high として、3 段階の評価点を推定するモデルを作成する。モデルごとの精度、macroF1 スコア、モデル平均適合率、モデル平均再現率を表 6 に示す。

表 6 3 段階評価推定実験の結果

モデル	精度	F1 スコア	適合率	再現率
立地	0.80	0.77	0.78	0.77
部屋	0.78	0.76	0.77	0.74
食事	0.79	0.74	0.75	0.75
風呂	0.68	0.70	0.69	0.72
サービス	0.78	0.75	0.76	0.74
設備・アメニティ	0.83	0.80	0.82	0.79
平均値	0.78	0.75	0.76	0.75

## 6. 考察・まとめ

表 6 で示した 3 段階評価の実験では、表 5 で示した 5 段階評価の実験に比べて、観点平均の精度・F1 スコアが共に約 20% 高くなっていることが読み取れることより、あるレビュー文 (ユーザーが付けた評価点が 5 であるレビュー文) に対して、評価点 4 か評価点 5 のどちらであるかは正確に分類できていないが、評価点 1-3 (low, middle) ではないが評価点 4 か 5 (high) のいずれかであるという推定はできていると考えられる。

これより、レビューテキストは書き手の主観であるため、評価点が異なっても内容は同じようなものであることが多いと考察できる。実際に、「A がよかった」という複数の同一レビュー内容でも、書き手 1 は評価点 5, 書き手 2 は評価点 4 と異なる評価点をつけている例が多数見受けられた。

また、レビュー文の内容の極性 (ポジティブ・ニュートラル・ネガティブ) を捉えて分類できていると考えられる。しかし、研究目的で述べた、SNS に投稿されているようなテキストのみの宿泊施設レビューに対して応用するには、5 段階の評価点の推定に関してより高い精度で行うことができる必要があると考えられる。

### 謝辞

本研究は JSPS 科研費 JP24K15193 の助成を受けたものです。

### 参考文献

- [1] “ホテルのクチコミ・お客様の声【楽天トラベル】” <https://travel.rakuten.co.jp/review/>
- [2] 張 博. レビューテキストの書き手の評価視点に対する評価点の推定. 言語処理学会 第 23 回年次大会 発表論文集. 2017, C5-5
- [3] “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/>
- [4] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto. Applying Conditional Random Fields to Japanese Morphological Analysis. Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. 2004, pp.230-237.
- [5] “Pretrained Japanese BERT models released/日本語 BERT モデル公開 - Tohoku NLP Group/東北大学 自然言語処理グループ”. <https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/>