

レビューの評価項目別自動スコアリングに基づく お出かけスポット推薦システムの提案

福嶋亮太郎[†] 中島 伸介[†]

[†] 京都産業大学 情報理工学部 〒603-8555 京都府京都市北区上賀茂本山

E-mail: †{g2054155,nakajima}@cc.kyoto-su.ac.jp

あらまし 観光サイトでは、観光地の場所や営業時間などの基本情報とともに、観光客のレビューや観光地に対する評価などが記載されている。ユーザは、評価やレビューを参考に観光地を選択することがあり、ユーザに与える影響は大きいと考えられる。本研究では、観光地レビュー文を分析し、詳細な評価項目ごとにスコアリングを行うことで、観光地の特徴を明確にする。また、評価項目別自動スコアリングに基づく観光地提案システムの構築を行うことで、個人の趣向に沿った観光地の提案を目的とする。

キーワード 情報推薦, レビュー分析, 観光スポット, 可視化

1. はじめに

じゃらん [1] やトリップアドバイザー [2] のように観光地を対象として評価やレビューを書き込める観光レビューサイト（以下、観光サイトと呼ぶ）がある。旅行を計画する際には、観光地レビューや観光地の評価を参考にして旅行先を決定する場合があるので、観光地レビューや評価は重要である。しかし、観光サイトの評価は1次元的に表現されており、わかりやすいという利点があるが、観光地の特徴やレビュー投稿者が観光地に満足した点が不明確という欠点がある。観光地の特徴やレビュー投稿者が観光地に対して満足した点を理解するためにはレビュー文を読む必要があるが、レビューを読むには時間と労力が掛かってしまい、行ってみたい観光地を発見するには多大な時間を要することがある。また、旅行者が未知の旅行先に訪れる際に、パッケージツアーやガイドブックのままに観光することがあるが、あらかじめ観光地やスケジュールを決定されていることが多く、個人の好みに合った観光地を訪れられない可能性がある。

そこで本研究では、観光サイトの「評価が一次元的で観光地の特徴が不明確」という課題を解決するために、観光サイトに投稿されたレビューを分析することで、評価項目別でスコアリングを行い観光地の特徴を明確にする。また、それらの結果を用いて、ユーザが観光地の評価項目に求める理想値を入力することで、個人の趣向に沿った観光地の推薦を行うシステムの開発を目的とする。

2. 関連研究

近年、商品の情報や観光地のレビューを掲載しているサイトが増加している。様々な商品をオンラインで販売している楽天市場 [3] や、コスメや美容に特化した総合情報サイトの@cosme [4] がある。また、旅行に特化したじゃらん [1] などの観光サイトもあり、多くの人々が利用している。レビューサイトが増加したことから、レビューを扱う研究が盛んに行われている。

矢部ら [5] は、コスメアイテム特徴の相違点を可視化し、相違点可視化システムを利用したコスメアイテム推薦手法を提案している。レビューを分析し、評価表現辞書を用いることでコスメアイテムの自動スコアリングを行い、購入済みアイテムと検討中のアイテムの相違点を可視化している。また、コスメアイテムの評価項目（潤い・浸透、色持ち、ナチュラル等）に理想値を入力することで、理想のアイテムを発見できるシステムを構築している。

住友ら [6] は、投稿された口コミ情報を収集し、感情分析することで各感情分析スコアを算出し、スコアに最も近いスポットを推薦している。しかし、このシステムでは周辺地域のスポットしか発見することができず、旅行先の観光地を発見することが出来ないため用途が非常に限定的である。

鳥山ら [7] は、年末年始に行われる無料サービスや夏に行われる祭りなど「季節特有の情報」と春には桜が咲き、秋は紅葉が綺麗など「季節ごとに変化する情報」を抽出するために、観光レビューからトピックを抽出し、トピックごとにクラスタリングを行うことで2種類の情報を抽出している。本論文では、季節で変化する情報ではなく、レビューの感情を分析する点で異なる。

また、観光写真から観光スポットを推薦する研究も行われている。中野ら [8] は、ユーザが観光地で撮影した写真を画像分析することで、写真の中からユーザ固有の好みタグを抽出することで、個人にパーソナライズされた観光スポット推薦を提案している。しかし、このシステムでは観光地で写真を撮ることを前提に作られており、必ずしも誰でも使用できるとは限らない。また、写真中から誤検出や異なるタグを抽出してしまっていることで、好みを抽出できているとは言い難い。

新井ら [9] は、Twitter から観光地でリアルタイムにつぶやかれたツイート情報を収集し、それらの情報を活用して観光ルートを推薦する手法を提案している。ユーザが出発地、到着地、出発時間、到着時間を条件として入力することで、入力条件を満たした観光スポットを提示するというメリットがあるが、必

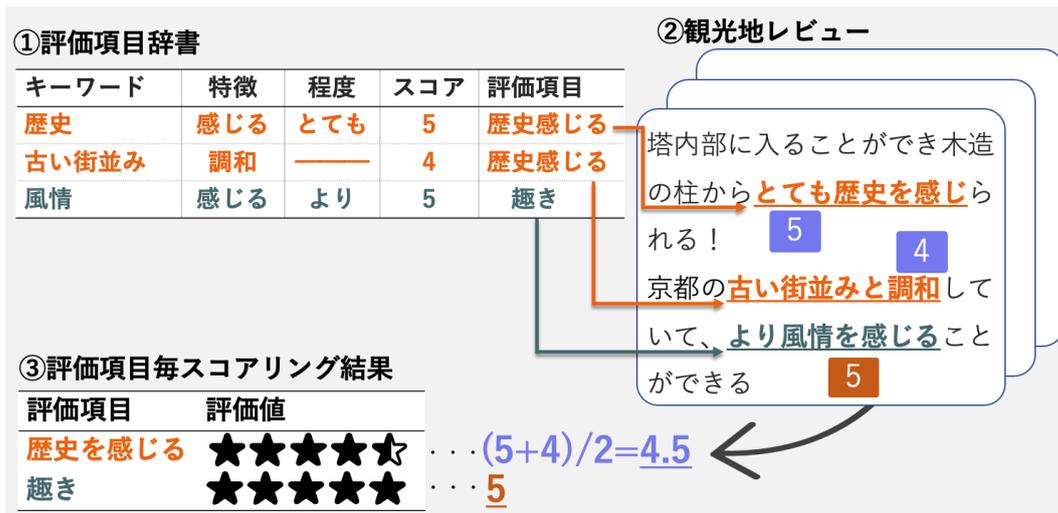


図1 スコアリング方式の例

ずしもユーザが行ってみたいと思える観光スポットを提示できるとは限らない。

3. レビューの評価項目別自動スコアリングに基づくお出かけスポット推薦システム

3.1 概要

観光地の情報を入手する手段として SNS, ガイドブック, 観光サイトなどが挙げられる。観光サイトは、観光地の評価やレビューを書き込めることから、多くのユーザが観光地を決定する際に参考している。しかし、観光サイトの評価は一次元で表現されており、レビュー投稿者が満足した点や観光地の特徴が不明確である。また、観光地の特徴やレビュー投稿者が満足した点を理解するには、レビューを読み進める必要があり、時間と労力が掛かってしまう。本研究のシステムでは、投稿者のレビューを分析することで観光地の特徴を詳細にスコアリングし、一次元的な評価から多次元的な評価を行う。また、ユーザが観光地に求める理想スコアを入力することで、最も理想に近い観光地を推薦するシステムの構築を目指す。

3.2 手動による評価表現辞書の作成と自動スコアリング

本説では、手動による評価表現辞書の生成と自動スコアリング手法について述べる。評価表現辞書には、評価項目、キーワード、程度がある。例えば、「古い街並み」「古都」などの評価項目に対応するキーワードがレビュー中に含まれている場合は、「歴史を感じる」という評価項目にスコアが付与される。キーワードの前に「とても」「非常に」などの程度語があれば、程度によってスコアが変化する。キーワード、程度は手動で評価表現辞書に追加する方法と 3.3 節で後述する自動で評価表現辞書に追加する方法がある。本稿で検討している評価項目は下記の通りである。

- 歴史を感じる
- SNS 映え
- ご飯が美味しい
- 落ち着く
- ホスピタリティ

- 四季折々
- 景色が綺麗
- デートスポット
- 京都っばい
- 癒される

今後も評価項目は追加する予定である。

次に、評価表現辞書を用いた各観光地のスコアリング方法を示す。図1を例にとると、観光地レビューに「古い街並みと調和」と書かれている。「古い街並み」という単語は「歴史を感じる」という評価項目に含まれるので、「歴史を感じる」という評価項目で4点のスコアが付与される。また、観光地レビューの中にある「とても歴史を感じ」という部分は5点のスコアが付与されているので、最終的に観光地に付与する「歴史を感じる」の評価項目スコアは、「古い街並みと調和」の4点と「とても歴史を感じる」の5点の平均値を取り、 $(5 + 4) / 2 = 4.5$ となる。

3.3 自動による評価表現辞書の作成手法

本説では、自動による評価表現辞書の作成方法について述べる。観光サイトでは多種多様な人がレビューを投稿しており、実際に使用しているキーワードを収集し、学習することで分類器モデルの構築を行う。

分類器モデルの構築手法

- 1 評価項目に該当するレビューを収集してラベル付けを行う
- 2 評価項目に該当しないレビューを収集する
- 3 1, 2を学習データとして分類器を生成する

構築した分類器に、観光地のレビューを読み込めせることで、評価項目に該当するキーワードを特定することができる。また、特定したキーワードを自動で評価表現辞書に追加することで、新たな言葉ができた場合や手動で評価表現辞書に追加し忘れていたキーワードを追加することができる。スコアリング手法については、3.2 節と同様に行う。

3.4 観光地推薦手法

本節では、評価項目別スコアリングを用いた観光地推薦手法について示す。ユーザは、観光地に求める条件をスライドバー

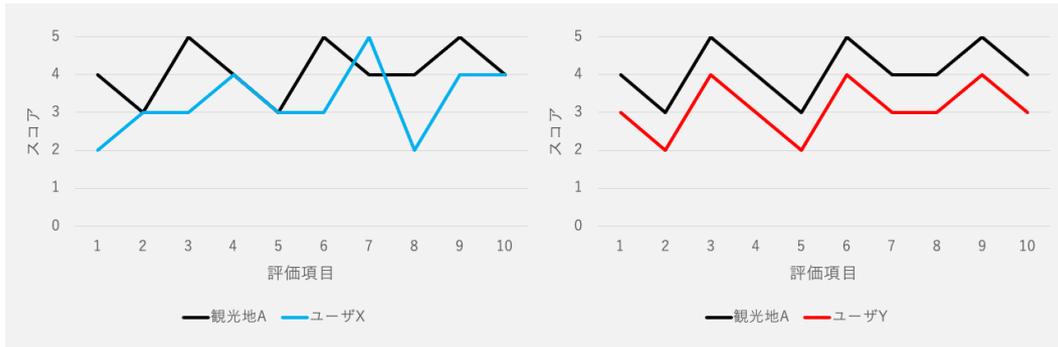


図2 ユーザの理想条件と観光地のスコアリング結果例

に入力し、観光地を分析して得られたスコアリング結果と最も近い結果を推薦する。このように、ユーザが尋ねてみたい観光地が明確になっていない場合でも、好みに沿った観光地を発見することができる。

このような理想条件を入力する際に、4, 5は入力しない人や1, 2は入力しない人がおり、ユーザによってゆらぎが生じることがある。ゆらぎを考慮して、ユーザの理想条件と観光地のスコアリング結果の類似度を計算することが重要である。次に、ユーザの理想条件と観光地のスコアリング結果の類似度を計算する手法を示す。

図2はユーザの理想条件と観光地のスコアリング結果の例である。左のグラフは、黒色の折れ線は観光地Aのスコアリング結果、青色の折れ線はユーザXが入力した理想条件である。同様に、右のグラフは、黒色の折れ線は観光地Aのスコアリング結果、赤色の折れ線はユーザYが入力した理想条件である。

表1 ユーザの理想条件と観光地のスコアリング結果の各データ

評価項目	観光地A	ユーザX	ユーザY
1	4	2	3
2	3	3	2
3	5	3	4
4	4	4	3
5	3	3	2
6	5	3	4
7	4	5	3
8	4	2	3
9	5	4	4
10	4	4	3

表1は図2の各データを表にまとめたものである。観光地Aのスコアリング結果とユーザXが入力した理想条件の差の総和は10で、観光地Aのスコアリング結果とユーザYが入力した理想条件の差の総和も10である。しかし、図2を見ると観光地AとユーザYのグラフの形は等しく、類似度が高いことがわかる。本研究では、評価項目とスコアの傾きを計算して類似度の計算を行う。また、傾きが等しい場合には、観光地のスコアリング結果とユーザが入力した理想条件との差を取り、最も差が少ない観光地を推薦する。

4. 評価実験

本稿では、提案システムの前段階として、評価項目の「歴史を感じる」のみでスコアリングした結果とじゃらんで「歴史」で検索した結果で評価実験を行う。

4.1 実験手法および目的

本システムでは、観光地の特徴を明確にし、個人の趣向に沿った観光地の提案を目的としている。そのため、従来システムと提案手法で比較検証を行い、本システムの有用性を検証する。

被験者に、従来システムで「歴史」と検索して表示された上位3件の観光地と評価項目の「歴史を感じる」でスコアリングした結果の上位3件の観光地を提示してアンケートに回答してもらう。従来手法で推薦された観光地は、「京都市学校歴史博物館」「歴史の道」「丹後ちりめん歴史館」である。また、提案手法で推薦された観光地は「北野天満宮」「えびす屋嵐山總本店」「京都国立博物館」である。提示された観光地を見て、図3のアンケートに回答してもらう。選択肢は「全くそう思わない」「そう思わない」「どちらでもない」「そう思う」「非常にそう思う」の5つから構成されており、単一選択で回答してもらう。

1	歴史を感じるランキングに相応しい
2	行ってみたいくない
3	この項目の上位にランキングされることが意外だ
4	従来の検索システムになさそう
5	モチベーションを高める

図3 評価実験のアンケート項目

5. まとめ

本稿では、観光サイトの評価が1次元で表現されており、観光地の特徴が不明確という課題を解決するべく、レビューを分析することで多次元的な評価を行う。また、複数の評価項目を用いてお出かけスポットを推薦することで、個人の趣向に沿ったお出かけスポット推薦システムを提案した。今後は、お出かけスポット推薦システムの構築に取り組んでいく。

謝 辞

本研究の一部は、科研費（課題番号：20H04293, 22H03700, 22K12281）および京都産業大学先端科学技術研究所（ヒューマン・マシン・データ共生科学研究センター）共同研究プロジェクト（M2001）の助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] じゃらん net, <https://www.jalan.net/kankou/>
- [2] Tripadvisor, <https://www.tripadvisor.jp/>
- [3] 楽天市場, <https://www.rakuten.co.jp/>
- [4] @cosme <https://www.cosme.net/>
- [5] 矢部沙也加, 上田真由美, 中島伸介, ”評価項目別スコアを用いたコスメアイテム間の相違点可視化システム”, DEIM Forum 2021, C14-2, 2021.
- [6] 住友千将, 石野拓也, 久保洗貴, 岳五一, ”口コミ情報に含まれる感情語に基づく類似スポット推薦システムの構築と実証実験”, パーソナルコンピュータ利用技術学会論文誌, 2020 年 14 巻 1 号 p.29-35, 2020.
- [7] 鳥山実桜, 灘本明代, ”観光レビューから季節特有な情報と季節によって変化する情報の抽出手法の提案”, DEIM Forum 2022, B43-1, 2022.
- [8] 中野広貴, 荒澤孔明, 渡邊稜平, 服部峻, ”観光写真から抽出した撮影者の好みに基づく観光スポット推薦”, IEICE Technical Report IN2018-80 (2019-01), 2019.
- [9] 新井晃平, 新妻弘崇, 太田学, ”Twitter を利用した観光ルート推薦の一手法”, DEIM Forum 2015, G7-6, 2015.