

「観光地戦略マップ」の可視化

- 口コミ・レビューデータを用いた観光地選好空間モデル -

鈴木英之*

Strategy Map of Tourist Destinations - Preference Space Modeling Using Online Reviews to Tourist Attractions -

Hideyuki SUZUKI*

Abstract: This paper acquired a distributed expression from review data described by tourist, and we made a positioning map for tourist spot-SNExt, the Kriging estimation was applied and the preference space model was created by adding the dimension of preference. It was confirmed that this "Strategy Map of Tourist Destinations" is useful for examining regional tourist spot strategies.

Keywords: 観光地経営 (destination management), クリギング (Kriging), 分散表現 (Distributed representation), t 分布型確率の近傍埋め込み法 (t-SNE), 多次元尺度構成法 (MDS)

1. 背景

観光地経営においてエビデンスに基づいた意思決定が適切に行われるために、大規模データや人工知能（機械学習）技法を利活用する動きが進展している。ただし、データに基づいて意思決定を行う組織能力や文化の醸成は民間企業においてすら未だ大きな課題である。複数のステークホルダーを緩やかに束ねた存在である観光地域もしくはDMOにおいて、戦略目標の策定や目指すべき方向性に関してコンセンサスを取り纏めることが、難題であることは容易に推察される。

2. 研究目的

これまでのところ、大規模データや、その取扱を容易にする人工知能技法は、観光マーケティング活動（SNS 活用形のプロモーション）や旅行者行動分析（観光地レコメンデーションや渋滞防止等の観光行動の最適化等）において既に多くの実用例が示されているものの、その多くは地域計画や地域戦略が扱う中長期的・戦略的な視座を持つものではなく、即時的なオペレーションや観光戦術的施策にかかるものである。

本報は近年の技術を活用した「観光地戦略マップ」を作成し、地域戦略策定に資することを目的とするものである。

3. 方法

本報は磯田（2009）および朝野（2010）の提案手法を土台とする。そこに大規模な口コミ・レビューデータを投入することで戦略マップの可視化を実現する。

3.1 非地理現象の地球統計学適用にかかる先行事例

磯田（2009）および朝野（2010）の先行研究は、ともに位置情報とは関わりのない分野の統計データに対して地球統計学的手法を適用することでデータの可視化を果たした。前者は、労働統計における職業間移動情報を二次元上の散布図に表現したうえで、そこに賃金水準や雇用増加率を3次元軸としてクリギングによる推定値を示し、「転職地図」を作成した。後者は、旅行地イメージにかかる質問紙調査データを二次元散布図に表現した上で、選好度をクリギング推定し、3次元ポジショニングマップとして示した。

* 正会員 合同会社ファイン・アナリシス (Fine Analysis, LLC.)

〒545-0021 大阪府大阪市阿倍野区阪南町 2-24-12 Tel : 06-7503-9426 E-mail : suzuki@finea.co.jp

3.2 レビューデータからの分散表現の獲得

本報は、先行研究と異なり調査データの代替として口コミ・レビューデータを利用する。tripadvisor.com に対する書き込みを対象に、2020年8月時点までに、52,584件の国内観光スポットに対する1,414,115件の日本語レビューを収集した。

収集データから散布図を作成するに際して、単語や文章を数値ベクトル化する機械学習手法である分散表現を利用する。本報では、文章に対する分散表現獲得アルゴリズムのひとつである Doc2Vec を用いて、各観光スポット別に、それぞれ書き込まれたレビュー文章群のベクトルを算出した (python ライブラリー gensim を用いて doc2vec pv-dm モデルを適用し 50 次元のベクトルを生成)。

3.3 ポジショニングマップの作成

ポジショニングマップは、分析スポットと比較対象スポットとの相対的位置関係を座標平面に表現した「地図」である。通例、多次元の調査データを主成分分析や多次元尺度構成法 (mds) を用いて縮約し二次元に描画する。

本報では、実務的用途から、広範囲のマクロ的視点と微細なミクロ的視点の両側面から2種類の可視化マップを作成した。

分析対象として、大阪市天王寺区に所在する天王寺動物園を選定した。全国に同一カテゴリー (動物園) 施設が一定数あり、かつ運営主体 (公営/民営) や利用者オケージョン (家族向け日帰りレジャー/目的体験型の連泊等) にバラエティが見込める事による。

マクロ視点のマップは、隣接する異なるジャンルのスポットとの比較を念頭に置いた上で、当該観光スポットの立ち位置を大胆かつ虚心坦懐にゼロベースから再検討する用途を想定した。近年、大規模多次元データの次元縮約で多用される t-SNE 法を利用し、数千箇所の観光スポットを分類して概観することを意図した。計算には Dmitry Ulyanov による Multicore-TSNE を利用した。レビュー数 50 件以上の全国の観光スポット中、天王寺動物園を含む、5,325 件を対象に 50 次元のベクトルを 2 次元空間に投影する。

ミクロ視点のマップは、分析対象と類似対照群との微細な比較を目的とする。主として同一カテゴリーの枠内でのリポジショニングの検討用途を想定した。天王寺動物園と類似する観光スポットとして cos 類似度の上位 100 スポットを抽出し、(1-cos 類似度) を距離として、100 スポット間の距離行列を作成し、これに多次元尺度構成法を適用する。

3.4 選好指標のクリギング推定とその表示

前項の2つのマップに対して、消費者による選好を表現する指標を地図上にクリギング推定することで選好空間モデルを作成する。選好度として用いたのは、tripadvisor に書き込まれた評価スコア (10~50 点) の平均値およびレビュー数である。またミクロ視点のマップでは動物園入園料 (入場料) を市場における条件付きの付加価値評価と見做した推定を行う。

4 結果と考察

4.1 レビューデータ分散表現の実用性

分散表現の獲得によって5万余の観光スポットにベクトルが付与された。表1は付与されたベクトルをもとに算出した類似観光スポットである。レビュー数が一桁程度のマイナーなスポットに対しても、ある程度に妥当なリスティングがなされているものとみられる。

短文表現のクチコミ・レビュー文では頻出表現句が多用される傾向にある。今回試行では、単純な単語レベルでの解析を行ったが、レビュー者が多用するフレーズ辞書を作成することでノイズの低減等が見込めるものと思われる。

4.2 ポジショニングマップ及び選好空間モデルの解釈可能性

t-SNE アルゴリズムによる次元縮約は、必ずしもマップ上の距離が類似度 (乖離度) と正確に比例しないことが指摘される。ただしマーケティング近視眼に陥ることなく市場の全体像を鳥瞰し、異なるカテゴリー間関係を理解するためには有用であると考えられる。図1散布図上では、動物園グループに隣接して牧場群、少し離れて植物園群やテーマパーク群が認められる。動物園群においては旭山動物園、隣接するテーマパーク群では東京ディズニーランド

(以下 TDL) 付近に満足度のホットスポットがみられる。

MDS で作成したマップ上には全国動物園の他、一部の水族館や、アンパンマンこどもミュージアム等の施設が含まれる(図 2)。これらスポットは子供連れの週末行楽のような類似のオケーション利用という共通点を持ちながらも、価格帯や体験志向(ふれあい体験やサファリ体験)、専門的ジャンル性(クマ、猿、アルパカ等)といった点で、ポジショニング上の多様性を持つものと見られる。

先に見た通り、旭山動物園、大牟田市動物園及びこども動物自然公園(埼玉県)等の施設を中心に高い満足度のエリアが示される。

満足度スコアがコストパフォーマンス評価に大きく左右されることは既往のマーケティング研究によって確認されている。そこで動物園入場料の等値線推定を行った。入場料が高い右上では満足度が低く、入場料が低い左下では満足度が高い一般的傾向が認められる。多くの公立動物園が所在する左下の象限で満足度が高いのは、利用者が受益する付加価値に見合った対価の一部を公費が負担することでコストパフォーマンスが上がっている可能性が考えられる。

4.3 ベクトル演算とリポジショニング

ポジショニングマップの主要な用法はリポジショニングの検討にある。ここではベンチマーキング先として、高評価スポットである TDL と旭山動物園を選定し、仮に天王寺動物園が TDL 化(旭山動物園化)するとした時の方向性や方策についての検討を試行する。方法として TDL(旭山動物園)に付与されたベクトルから天王寺動物園のベクトルを減算したベクトルをリポジショニングにかかるベクトルとし、これに近いベクトルを持つ単語群や観光スポットを解釈することでリポジショニング方策の検討材料とする(表 2)。

TDL 化に必要なイメージとして観光スポットとしては「レインボーブリッジ」や「横浜ベイブリッジ」、単語としては「プロポーズ」や「恋」、「恋愛」が挙がる。天王寺動物園と比較して TDL ではデートコースとしての利用シーンに強いことが想像される。一方、旭山動物園化に必要なのは、「大雪山」や「ネイ

チャークルーズ」等、大自然のスケール感である。ベクトルの近い単語として挙がるのも、主として「大潮」、「流氷」、「雲海」等の自然現象である。

TDL 化と旭山動物園化とは方向性が全く異なるものの、両者のリストに共通して「長瀬ラインくんだり(全長 6km:40 分のダイナミックな自然を満喫する豪快なアクティビティ)」が入っている点も興味深い。

5 まとめ

今回の試行で、これまで限定的な質問紙調査や、内部によるワークショップのみによって検討されてきた観光地戦略の検討に際して、口コミ・レビューデータの情報を活用した当該戦略マップの利用可能性について示すことができたものと考えている。

参考文献

- 磯田弦(2009)「転職地図」を使った職業別労働市場の分析 ―地理情報システムの非地理空間への応用―, 地理情報システム学会学術講演論文集, 18, CD-ROM.
- 朝野熙彦(2010)選好空間モデリングへの地球統計学の応用, 日本行動計量学会第 38 回大会抄録集 pp. 282-283
- 野守耕爾, 神津友武(2014)「観光クチコミデータを用いた類似観光地の発見と満足形成要素の分析」, サービス学会第 2 回国内大会講演論文集, pp. 95-100
- 愛媛県(2017)「ビッグデータを活用した観光戦略等検討委託事業報告書」
- 北海道観光振興機構(2017)「平成 28 年度北海道外国人観光客再訪促進事業調査事業報告書」
- 観光庁(2020)「AI(人工知能)等導入による旅行サービスの高度化事業調査報告書」Q. Le, T. Mikolov, 2014. Distributed Representations of Sentences and Documents

表1 観光スポットとその類似スポット

天王寺動物園 (件数: 347, 単語数: 49,409)				寺西家阿倍野長屋 (件数: 8, 単語数: 839)			
観光スポット名	cos類似度	件数	単語数	観光スポット名	cos類似度	件数	単語数
那須ワールドモンキーパーク	0.864	71	10,363	換断屋敷	0.936	3	441
平川動物公園	0.862	113	16,403	まちや館	0.916	4	370
京都市動物園	0.850	144	22,369	湖北観光情報茶屋四居家	0.915	5	566
長崎バイオパーク	0.848	136	22,484	ハルナグラス	0.914	7	991
神戸どうぶつ王国	0.843	421	64,477	四日市市歴史民俗資料館	0.914	8	1,079
福岡市動物園	0.836	148	21,050	倉吉レトロ館	0.912	5	532
群馬サファリパーク	0.835	142	24,468	胆沢城跡	0.911	5	537
静岡市立日本平動物園	0.835	216	31,900	デ・モレーンいまね	0.909	6	1,007
東山動物園	0.835	604	90,148	木戸跡広場	0.908	9	859
姫路市立動物園	0.834	127	18,305	ギャラリー再会	0.908	5	766
よこはま動物園ズラシア	0.833	453	71,013	中山道鶴沼宿 脇本陣	0.907	4	457
大牟田市動物園	0.832	40	6,099	旧南部鎗造所 キュボラ	0.907	2	212
札幌市山手動物園	0.829	589	87,679	坂田医院旧診療所 建築物	0.907	10	1,272
いしかわ動物園	0.828	56	9,407	脇本陣滝沢屋	0.907	4	523
とくしま動物園	0.823	24	3,315	下田藍染工房	0.907	2	223
熊本市動物園	0.822	91	13,126	吉田家住宅	0.907	5	735
金沢動物園	0.820	102	17,215	小城市北園街道ほんまち町歴史館	0.906	12	1,364
釧路市動物園	0.816	85	13,278	木屋本店	0.906	2	331
千葉市動物公園ふれあい動物の里	0.815	210	30,097	高梁市郷土資料館	0.905	12	1,508
池田動物園	0.812	62	8,684	旧笠石家住宅	0.905	4	389

表2 リポジショニングにかかるベクトルの類似スポット

【東京ディズニーランド】→【天王寺動物園】				【旭山動物園】→【天王寺動物園】			
観光スポット名	cos類似度	件数	単語数	観光スポット名	cos類似度	件数	単語数
東京ディズニーシー	0.656	2,279	340,267	雪の大谷	0.538	160	23,711
東京ディズニーランド	0.642	2,507	378,469	知床ネイチャーコース 日帰りツアー	0.529	49	8,978
レインボーブリッジ	0.570	778	88,524	三四郎島のトンボロ	0.522	29	4,574
長瀬ラインくんだり	0.564	154	24,203	旭山動物園	0.520	1,961	316,632
横浜ベイブリッジ	0.552	296	35,008	厳美溪	0.515	204	26,451
ユニバーサル・スタジオ・ジャパン	0.530	3,481	551,092	知床横断道路	0.515	76	9,482
ニセコレジャビュア	0.518	48	8,941	千景敷カール	0.508	261	40,622
クイーンさまみ	0.510	27	3,749	長瀬ラインくんだり	0.505	154	24,203
伊良部島	0.508	381	51,381	清水渓流広場 清瀬の滝・亀岩の洞窟	0.504	235	33,824
小笠原海運	0.506	24	4,353	抱返り溪谷	0.502	128	17,148
野沢温泉スキー場	0.501	249	38,476	白ひげの滝	0.500	285	33,776
青函フェリー	0.496	16	2,608	香嵐溪	0.498	281	42,268
戎橋	0.495	250	28,012	遠音別川	0.495	37	5,743
生口橋	0.494	49	5,621	立山黒部アルペンルート	0.494	463	72,888
中洲歓楽街	0.491	572	67,480	百合ヶ浜	0.494	67	9,062
清洲橋	0.488	45	5,023	長岩岩壘	0.493	249	30,370
隅田川	0.485	185	24,453	青の洞窟	0.491	31	4,630
フェリーさんふらわあ	0.484	37	5,016	知床・伊達タンチョウフェスティバル	0.488	63	8,748
すずきの	0.483	1,280	152,376	十二湖	0.487	250	37,272
角島大橋	0.482	797	99,975	父母ヶ浜	0.484	37	5,545

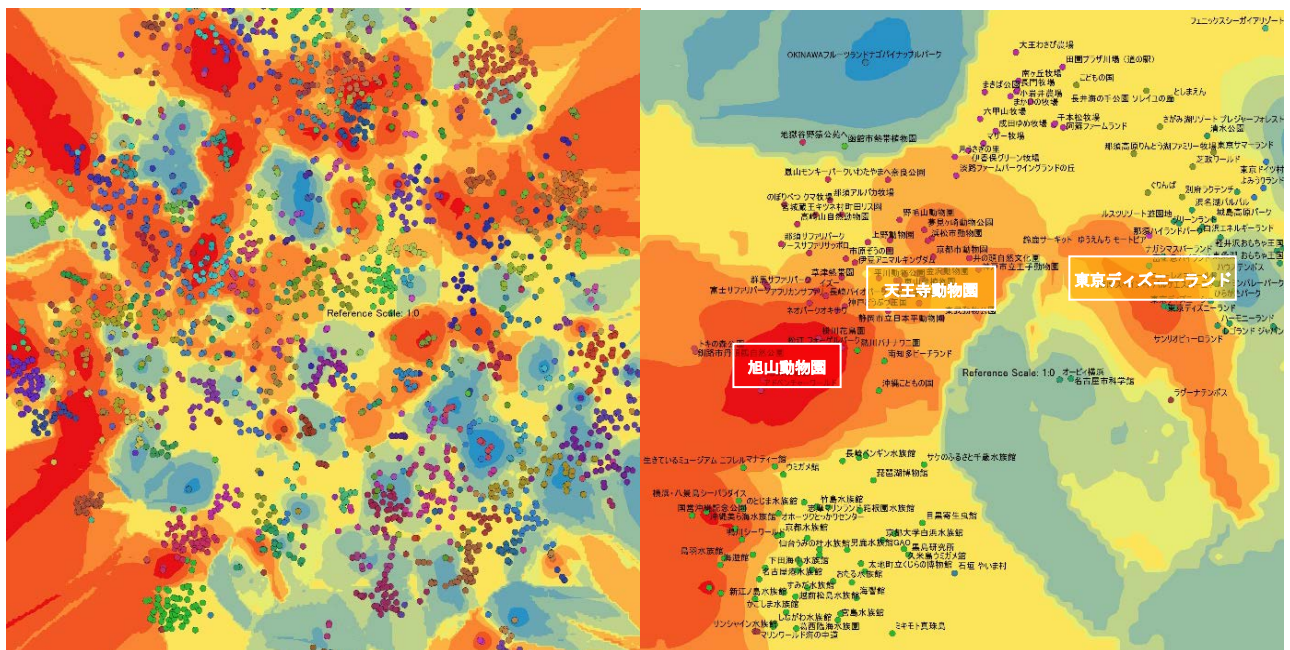


図1 t-SNEによるマクロ視点の「観光地戦略マップ」(左)全体図 (右) 動物園クラスター付近の拡大図

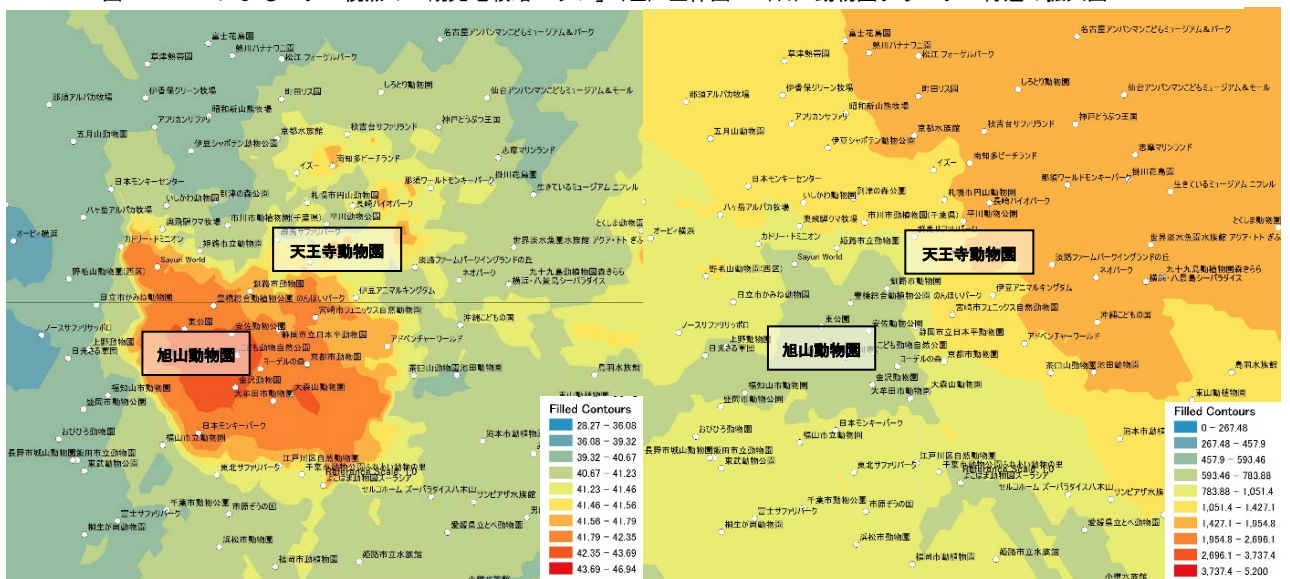


図2 MDSによるミクロ視点の「観光地戦略マップ」(左) 評価スコア (右) 入場料 (いずれも暖色が高い値を示す)