

娯楽施設に対する期待度が満足度に影響する要因について —XAI を用いた大規模アンケートのデータ解析—

長谷川 康邦[†] 齊藤 史哲[†]

[†]千葉工業大学 先進工学部 〒275-0016 千葉県習志野市津田沼 2-17-1

E-mail: s19c3100bk@s.chibakoudai.jp, fumiaki.saitoh@p.chibakoudai.jp

あらまし 娯楽施設は余暇時間の充実や心身の休息などの役割を担っているが、近年では利用者の減少や閉館などの課題に直面している。娯楽施設を維持するためには顧客満足度の向上が手段の一つとして挙げられ、これによりサービス改善や発展につながった事例はいくつかある。顧客満足度には複数の種類の期待度と満足度の相対関係が成り立つが、先行研究では単一項目のみで満足度要素の影響を調査しており、複数の性質を考慮した調査が不十分である。そこで、本研究ではオリコンデータセットに対して XAI 手法である SHAP と Boruta を適用することで、複数の期待度や期待度と満足度との差が総合満足度に対する影響の及ぼし方を解析し、重要な顧客期待要素を明らかにした。

キーワード 娯楽、娯楽施設、顧客満足度、XAI、Boruta、SHAP

1. はじめに

1.1. 娯楽施設の課題

娯楽施設は余暇時間の充実や心身の休息などの役割を担っているが、近年では利用者の減少や閉館などの課題に直面している。例えば、映画館は 1958 年をピークに来場者数が減少し、低水準を維持している（図 1）。これはメディアの発展により映画がテレビ放送されたり、映像配信サービスにおいて映画が配信されたりするようになったことで、映画館に行かなくとも映画体験を得られるようになったことが大きな要因とされている。また、テーマパークは施設の閉園が問題となっている。具体的には 2017 年のスペースワールド、2020 年のとしまえんの閉園などが挙げられる。考えられる要因の一つにレジャーの多様化が挙げられ、テーマパーク以外の娯楽施設で充実した余暇時間を過ごすことができるようになったことが考えられる。

1.2. 娯楽施設と顧客満足度

娯楽施設を維持、発展させるための手段の一つとして顧客満足度の向上が考えられる。顧客満足度は製品やサービスに対する利用者の評価尺度の一つで、期待度と満足度との相対関係が成り立つ。また、顧客満足度の向上により利用者増加やサービスの質の改善などの効果が現れた事例はいくつかある。

顧客期待要素には複数のそれぞれ異なる性質をもつもので構成され[1]、複数の種類の期待度が満足度に何かしらの影響を与えることから、複数の期待度の中で特に重要な期待度を明らかにすることが重要であると考える。

2. 先行研究

2.1. 期待構成要素と娯楽施設

河田（2019）[2]は、観光地・観光施設のブランドイメージに対する評価への、顧客満足度要素の影響を明らかにすべく、ユーザーの過去から現在、そして将来における期待を含めた尺度を用いて、顧客期待要素による顧客満足度への影響を調査している。調査の結果、観光地や観光施設でのブランドイメージの向上は、利用者の顧客満足度や顧客ロイヤルティに良い影響を与えることが明らかとなっている。しかし、調査対象となつた観光施設や観光地の種類が複数存在しているため、全体的な傾向を明らかにすることは可能であるが、個々の施設でそれぞれ抱く特有の期待度が満足度に及ぼす影響や、複数の種類の期待度が満足度に及ぼす影響の把握が難しい。

2.2. リピーターについて

近年の娯楽施設ではリピーターの確保が課題となっており、リピーターを確保できるかどうかが娯楽施設の維持に影響する要素の一つであると考えられる。小林ら（2005）[3]はリピーターと初回利用者それぞれの特性の把握が重要であると述べている。

佐藤ら（2011）[4]は観光地へリピーター確保の方策を考察すべく、リピート行為を行う層（リピーター）とそうでない層（ノンリピーター）の違いや、リピーター特有の行動特性や意識について、関東 1 都 6 県を対象とした Web 調査と、箱根や熱海を訪れた人を対象とした入込調査を行っている。Web 調査ではリピーターの再訪行動を促す要因として「習慣性」や「自動車保有の有無」が挙げられている。また、入込調査

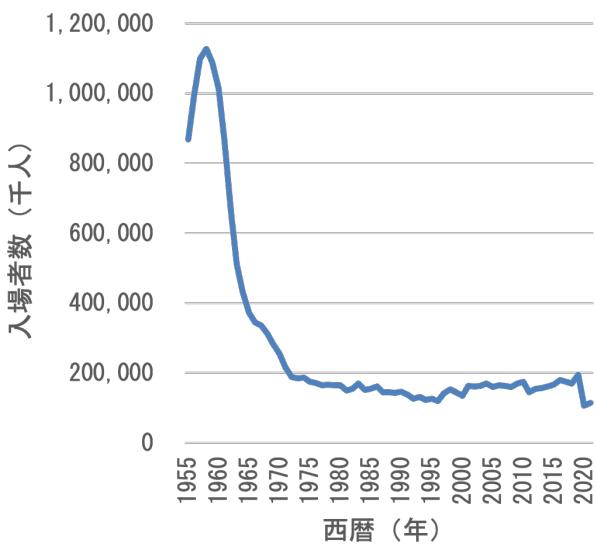


図 1 映画館の年代別の入場者数の推移。
(日本映画産業統計より)

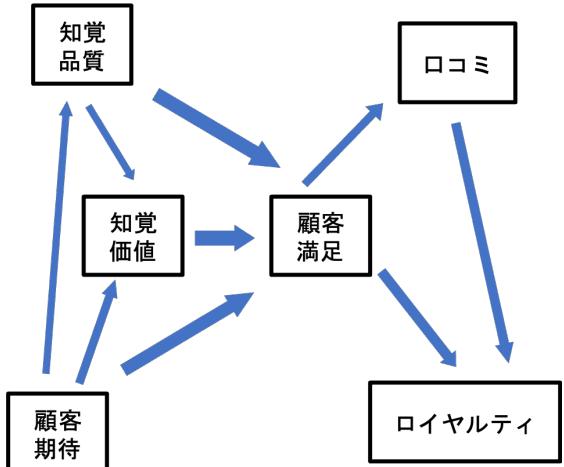


図 3 JCSI の因果モデル

及び満足度と期待度の差を算出し、5 章で述べる提案手法を用いて総合満足度に寄与する顧客満足度要素を明らかにする。

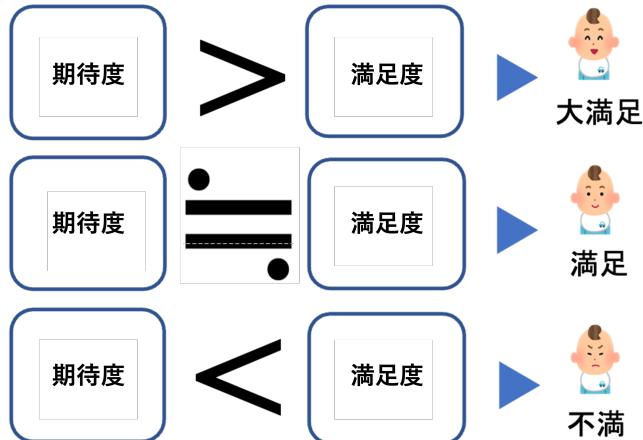


図 2 期待度と満足度の相対関係

では、箱根ではリピーターの好むパターンが明らかにできた一方で、熱海ではリピーターとノンリピーターとの差異が見られず、観光地を十分に周遊できていないことを指摘している。

リピーターの特性に関する研究は、全体的な傾向や個人の属性の違い、観光地域ごとの研究が幾つか存在するが、個々の施設を対象とした研究は筆者の知る限りでは存在しない。

3. 研究目的

本研究では娯楽施設の維持のために重要な顧客満足度要素を明らかにすることを目的とする。具体的には、映画館とテーマパークのアンケートデータを用い、それらのデータの満足度と期待度の数値の抽出、

4. 分析データ

今回使用したデータ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、株式会社 oricon ME より提供されたオリコンデータセット[5]のうち「映画館」と「テーマパーク」のアンケートデータである。映画館の調査期間は 2022 年 1 月 8 日～23 日で、テーマパークは 2020 年 1 月 6 日～14 日である。有効回答数は映画館が 7442 件、テーマパークが 5454 件である。

5. 分析手法と分析設定

5.1. Boruta

Boruta は仮設検定とランダムフォレストなどのアンサンブル学習を組み合わせた教師あり学習で、重要な変数（特徴量）を選別、抽出することができる手法である。教師あり学習では変数重要度（重回帰分析では偏回帰係数、ランダムフォレストではジニ不純度など）を参照することで、どの変数が重要かどうかを判断できるが、重要であるかどうかの基準が存在しないために重要度の大小比較のみでの判断となるため、真に重要な変数を見落とす恐れがある。Boruta は入力された変数のコピー（偽の特徴量）と一緒に学習させた時の精度低下を利用し、仮設検定を用いて重要な変数の基準を設けることで、真に重要な変数を過不足なく抽出できるようにした手法である。

表 1 映画館の回答者の内訳（地域別）

地域	回答者数（人）
北海道	340
東北	489
関東	2879
甲信越・北陸	493
東海	766
近畿	1206
中国・四国	608
九州・沖縄	661
合計	7442

表 2 映画館の回答者の内訳（年代別）

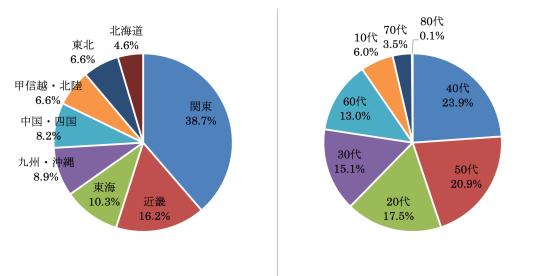
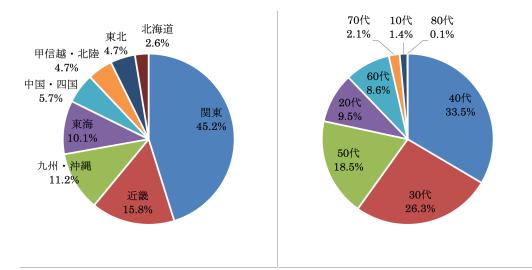
年代	回答者（人）
10代	444
20代	1301
30代	1122
40代	1778
50代	1559
60代	969
70代	258
80代	11
合計	7442

表 3 テーマパークの回答者の内訳（地域別）

地域	回答者数（人）
北海道	141
東北	255
関東	2464
甲信越・北陸	259
東海	553
近畿	860
中国・四国	310
九州・沖縄	612
合計	5454

表 4 テーマパークの回答者の内訳（年代別）

年代	回答者（人）
10代	79
20代	517
30代	1437
40代	1828
50代	1010
60代	468
70代	112
80代	3
合計	5454

図 4 映画館の回答者の割合
(左 : 地域別, 右 : 年代別)図 5 テーマパークの回答者の割合
(左 : 地域別, 右 : 年代別)

Boruta は以下のステップで変数を選択する。

- Step.1: 入力された変数（特徴量）をコピーし、シャッフルしたものを偽の特徴量（Shadow Importance）として入力データに組み込む。
- Step.2: 偽の特徴量を加えたデータを学習させ、変数重要度を算出する。
- Step.3: 偽の特徴量のうち最も大きい変数の特徴量より大きい入力データの特徴量を記録する。これを複数回実施する。
- Step.4: 仮設検定を用いて重要な変数を選択する。

5.2. SHAP

SHAP は各特徴量の貢献度を算出する方法を用いて、ブラックボックスな学習モデルを説明する方法として用いられる XAI 技術である。ランダムフォレストやニューラルネットなどの非線形な学習手法は、線形回帰と比べて高い精度の学習モデルを生成できるが、どの特徴量がどのように学習モデルに貢献したかが判断できないという欠点がある。SHAP は貢献度を計算する手法を用いて、入非線形な学習モデルに解釈性を付与することができる。これにより、学習モデルの全体的な傾向を把握できる他、入力された特徴量の分布に応じた分析が可能となる。

モデルの予測値を $f(x)$ としたとき、このモデルの貢献度は式 1 のように表される。

$$f(x) = \phi_0 + \sum_{j=1}^n \phi_j \quad (1)$$

ここでここで ϕ_0 は予測値の期待値を表す. ϕ_j は式 2 のように表される.

$$\phi_j = \frac{1}{|\mathcal{J}|!} \sum_{S \subseteq \mathcal{J} \setminus \{j\}} (|S|! (|\mathcal{J}| - |S| - 1)!)(v(S \cup \{j\}) - v(S)) \quad (2)$$

ここで \mathcal{J} は特徴量の集合で, $|\mathcal{J}|$ は特徴量の総数, $|\mathcal{J}|!$ は特徴量の組み合わせの総数を表す. S はある特徴量 j を差し引いたときのすべての組み合わせを表す. $v(S \cup \{j\}) - v(S)$ はある特徴量 j の値がわかっているときとわかっていないときの予測値の差分である.

式 1 より, 予測モデルの予測値を式 2 によって算出された貢献度に分解することで, 予測モデルの寄与度を算出する.

5.3. 分析設定

本研究では「総合満足度」を目的変数とし, 「期待度」「期待度と満足度の差」を説明変数として, 以下の手順でデータ分析を行なった.

Step.1: 対応する期待度と満足度を 1 対 1 で対応させるため, 対応する満足度の数値の平均を取る.

Step.2: 対応する期待度と満足度の差を取る.

Step.3: 目的変数を「総合満足度」とし, 説明変数を項目別の「期待度」「期待度と満足度の差」とし, そのデータを分析に用いる.

Step.4: Boruta を用いて重要な特徴量を抽出する.

Step.5: SHAP を用いて特徴量の貢献度を計算する.

※Step.4 及び Step.5 では施設ごとに分析

本研究で用いた Boruta のパラメータは, 最大施行回数を 100 回, 有意水準を 5%とした. Boruta 及び SHAP の変数重要度の算出には XGBoost (勾配ブースティングを用いたアンサンブル学習) を用い, その反復回数は 5 回とした.

今回は映画館 3 施設, テーマパーク 3 施設の計 6 施設を分析対象とした. 分析の際は施設ごとに, リピーターと初回利用者にデータを分けて分析を行った. 分析データの内訳は表 5 の通りである.

6. 分析結果と考察

6.1. 分析結果

図 6.1～図 6.12 に分析結果の図を載せる. それぞれの左図はリピーターの回答に対するもので, 右図は初利用者の回答に対するものである. これらの縦軸は変

数のインデックスであり, 横軸は計算されたそれぞれの指標値に対応している (Boruta は右に触れると重要度が高いことを表す. SHAP は真ん中の軸を中心に左に触れると総合満足度を低くさせ, 右に触れると総合満足度を高める影響を与えることを表す). Boruta の箱ひげ図の色は偽の特徴量と比較した仮設検定の結果に対応しており, 重要要素かそうでないかを表している.

(緑: 重要要素, 赤: 重要でない要素, 青: 偽の特徴量). SHAP の色は変数の大きさ (紫: 期待していた or 期待以上, 黄色: 期待していなかった or 期待はずれ) を表している.

6.2. 考察

映画館及びテーマパークで共通している点は「チケットの買いやすさ」の重要度が高いことである. チケット購入は施設での体験を得る前の段階で印象に残る要素であり, この要素が施設の第一印象を左右すると考えられる. ただし, 今日ではチケットは紙のチケットや電子チケットの種類が存在し, 購入方法も当日に現地で購入することができたり, あらかじめインターネットで予約購入することもできたりする. そのため, 購入したチケットの種類によって印象に差が出る可能性がある.

「期待度」と「期待度と満足度の差」の両方の要素が重要であると算出された要素は, 対象施設のマーケティングや宣伝内容, 利用者の口コミなどから得られる情報 (利用者の先入観, または固定観念) を表している可能性があると考える.

初回利用者のみで算出された要素として, 映画館 B では「利用のしやすさの期待度と満足度の差」, テーマパーク A では「ショップの充実さの期待度と満足度の差」が挙げられる. これらの SHAP の分析結果では, 期待以上だった回答者, または期待はずれだった回答者のどちらも分布しており, 期待以上であれば総合満足度を上げ, 期待はずれであれば総合満足度を下げる影響を与えていることが読み取れる. なお, 具体的に何に期待していたのか, 何に対して期待はずれだったのかといった内容はアンケートデータの項目だけでは読み取ることが難しいので, 深掘りするためには追加で調査をする必要があると考える. また, 初回利用者の分析結果はサンプル数の影響 (サンプル数が少ない可能性) を受けていると考えられ, 分析結果から読み取れる内容の傾向に偏りがあると考えられる.

映画館では「上映作品の充実さ」の重要度が共通して高かった. 映画館を利用する目的として考えられるものは「気になる作品を観ること」であり, 観たい作

表 5 分析対象の施設の内訳

施設名	リピーター	初回利用者	合計
映画館 A	1046	81	1127
映画館 B	1156	69	1225
映画館 C	745	68	813
テーマパーク A	238	105	343
テーマパーク B	433	161	594
テーマパーク C	676	85	761

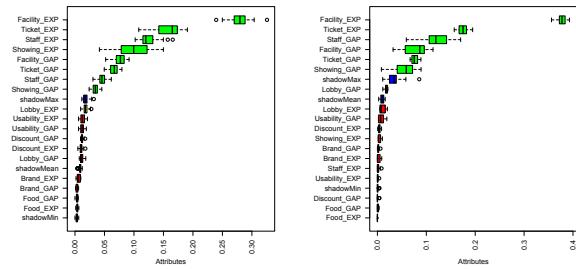


図 6.1 映画館 A の Boruta の分析結果
(左 : リピーター, 右 : 初回利用者)

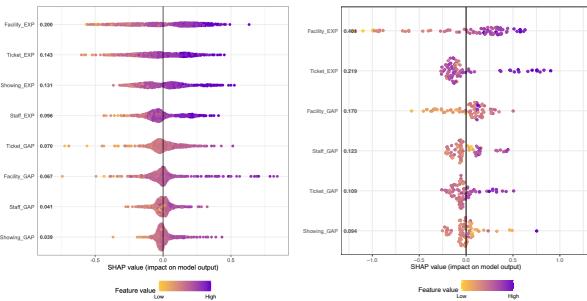


図 6.2 映画館 A の SHAP の分析結果
(左 : リピーター, 右 : 初回利用者)

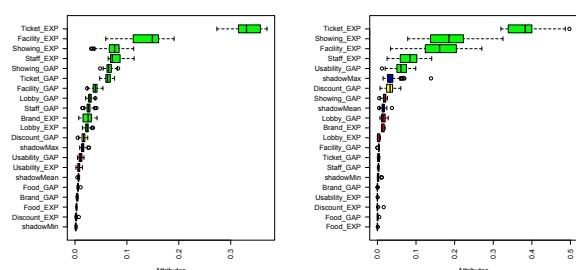


図 6.3 映画館 B の Boruta の分析結果
(左 : リピーター, 右 : 初回利用者)

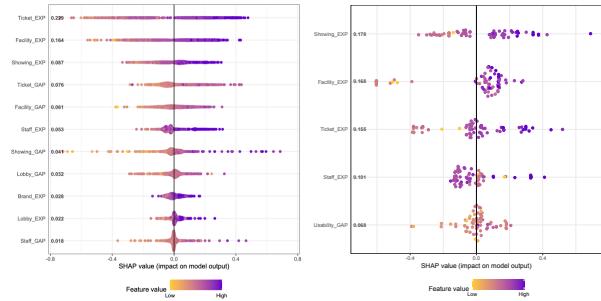


図 6.4 映画館 B の SHAP の分析結果
(左 : リピーター, 右 : 初回利用者)

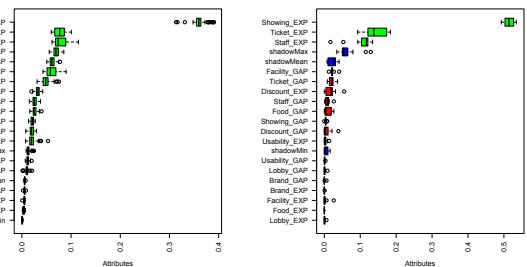


図 6.5 映画館 C の Boruta の分析結果
(左 : リピーター, 右 : 初回利用者)

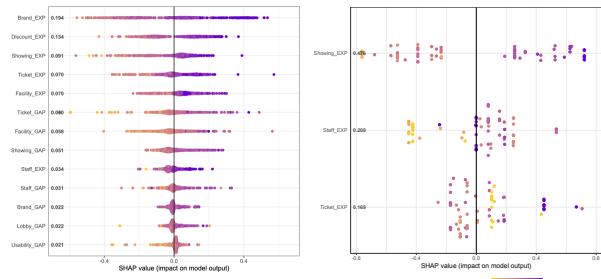


図 6.6 映画館 C の SHAP の分析結果
(左 : リピーター, 右 : 初回利用者)

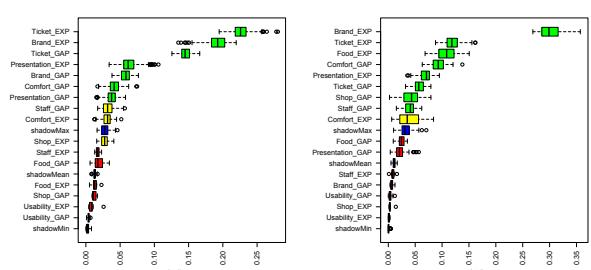


図 6.7 テーマパーク A の Boruta の分析結果
(左 : リピーター, 右 : 初回利用者)

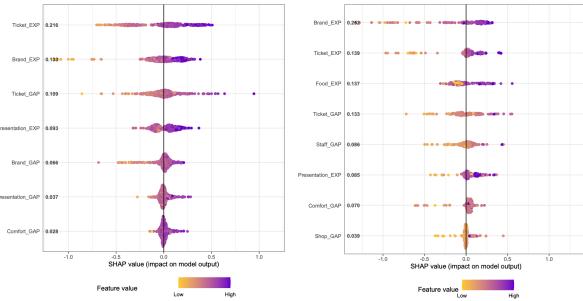


図 6.8 テーマパーク A の SHAP の分析結果
(左 : リピーター, 右 : 初回利用者)

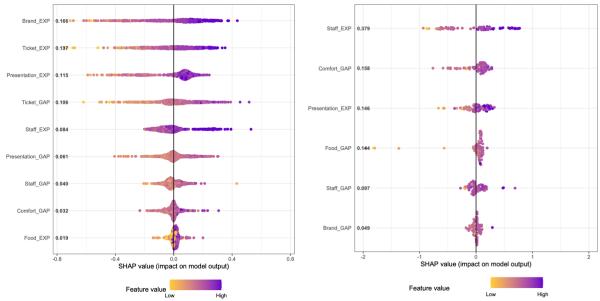


図 6.12 テーマパーク C の
SHAP の分析結果
(左 : リピーター, 右 : 初回利用者)

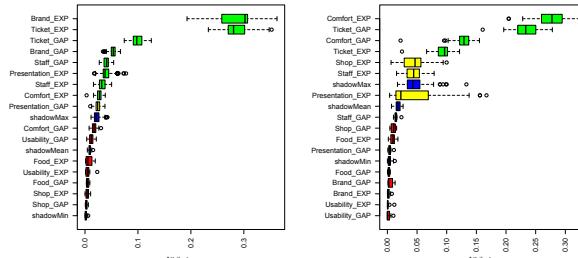


図 6.9 テーマパーク B の
Boruta の分析結果
(左 : リピーター, 右 : 初回利用者)

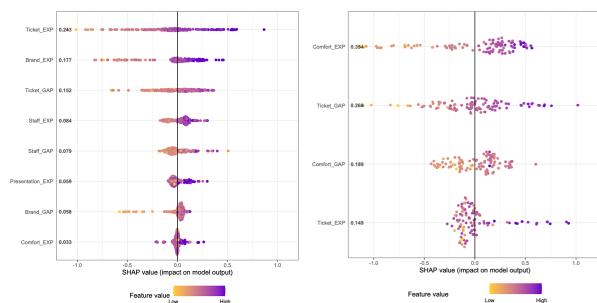


図 6.10 テーマパーク B の
SHAP の分析結果
(左 : リピーター, 右 : 初回利用者)

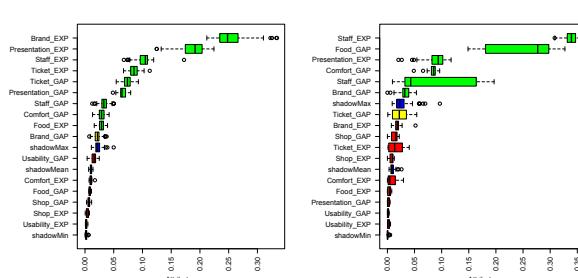


図 6.11 テーマパーク C の
Boruta の分析結果
(左 : リピーター, 右 : 初回利用者)

品が上映されているかで満足度を左右するものと考えられる。施設ごとではリピーターと初回利用者とともに、映画館 A は「設備の充実さの期待度」、映画館 B は「チケットの買いやすさの期待度」、映画館 C は「ブランドイメージの期待度」の重要度が最も高かった。

テーマパークでは特にリピーターで「ブランドイメージ」の重要度が高かった。テーマパークはその施設で 1 つのブランドを形成する要素であると考えられるため、施設で得た体験と体験前のイメージが満足度を左右するものと考えられる。

施設ごとで最も高かった重要な要素は、テーマパーク A のリピーターでは「チケットの買いやすさの期待度」、初回利用者は「ブランドイメージの期待度」の重要度が高かった。テーマパーク B では、リピーターは「ブランドイメージの期待度」、初回利用者は「パーク内の快適さの期待度」、またテーマパーク C では、リピーターが「ブランドイメージの期待度」、初回利用者は「スタッフの対応の期待度」の重要度が最も高かった。

今回の分析では Boruta と SHAP を用いたが、Boruta を用いたことで変数重要度の大小比較だけでは判断できなかった重要な要素を過不足なく判別できるようになった。変数重要度の大きさだけでは回答者が期待以上だったか期待はずれだったかを判断できなかったが、SHAP を用いることで回答者の分布から期待以上だったか期待はずれだったかを判断できるようになった。娯楽施設の利用前のイメージと利用後のイメージのギャップを探ることは、現在の宣伝の仕方でどのようなイメージを得ることができるのか、今後どのような宣伝が効果的なのかを考案する材料になり得ると考えた。

7. まとめ

本研究ではオリコンデータセットの「映画館」と「テーマパーク」のアンケートデータを、Boruta と SHAP の分析手法を用いて重要な顧客期待要素を明らかにした。これらの分析手法を用いることで重要な顧客期待要素と、その影響の与え方、回答者の分布と顧客期待

要素の関連性（どの程度の期待だったか、それが総合満足度にどのように影響を与えるか）を明らかにすることができた。施設ごとに分析することで全体的な傾向だけでなく、その施設特有の重要顧客期待要素や、利用者へのマーケティング効果と利用後の施設の印象を考察することができた。今後は別の項目別の期待度と満足度の差の算出方法について別の手法の検討や、別の分析手法を用いた分析、さらに、変数間の相関や交互作用を考慮した分析など、娯楽施設に対する期待度と満足度についてのさらなる分析や別の視点の考察などを進めたい。

謝辞

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を通じて株式会社 oricon ME 様よりデータをご提供いただきました。この場合をお借りして厚く御礼申し上げます。また、本研究は科学的研究費（基盤 C) 19K04887 による支援を受けたものです。ここに記して感謝いたします。

参考文献

- [1] 古屋 繁, 櫻井 貴章, 羽渕 拓哉, 杉山 和雄, “顧客満足度における「期待」マネジメントのための構造化”, 日本デザイン学会 デザイン学研究, 2014.
- [2] 河田 浩昭, “観光地・観光施設に対するブランドイメージ評価の刊行者ロイヤルティ形成への影響に関する考察”, 日本国際観光学会論文集(第26号), 2019.
- [3] 小林克行, 岩倉成志, “観光交通を対象としたトライアル・リピート需要の予測方法に関する研究”, 土木計画学研究・講演集, 2005.
- [4] 佐藤 友里子, 岡本 直久, "国内旅行におけるリピーターの行動特性及び釀成要因に関する研究", 土木学会論文集 D3 (土木計画学), Vol. 67, No.5 (土木計画学研究・論文集第 28 卷), 1_455-1_564, 2011.
- [5] 株式会社 oricon ME, “顧客満足度調査データ”, 国立情報学研究所情報学研究データリポジトリ(データセット), 2019, <https://doi.org/10.32130/idr.10.1>, (参照: 2022-07-11)
- [6] Molnar, Christoph, “Interpretable machine learning. A Guide for Making Black Box Models Explainable”, 2019. <https://christophm.github.io/interpretable-ml-book/>. (参照: 2022-10-10)