# S08 評価傾向の特徴量に基づいたヘアサロンにおける顧客プロファイリング

IDRユーザフォーラム2025　橋本和果，　齊藤史哲（青山学院大学）

## 研究の背景

- 美容院業界は店舗数の増加が続くオーバーストア状態
- 画一的なマーケティング施策が多く、顧客の多様なニーズに応えきれていない
- 人気店や人気スタイリストに予約が集中し店舗やスタッフが埋もれてしまう

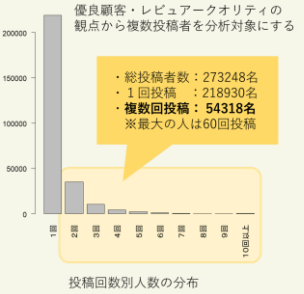
### 周辺研究について

- ヘアサロンを対象としたレビューデータや評価データを解析する研究が近年盛んに行われている[1-3]
- 評価者によって評価の仕方の傾向が異なるが、従来研究ではその点に関する議論が十分になされていない
- 顧客の特性にあった施策の提供、パーソナライズド推薦などのニーズ

顧客の行動や評価傾向に基づいた特性の定量化と分類が求められる

### 使用したデータ

- リクルート様からご提供いただいたHot Pepper Beautyのクチコミデータ
- 各評価者のデモグラフィクス：性別，世代
- 各店舗に対する5段階評価：
  - ムード，
  - サービス，
  - テクニック，
  - メニュー，
  - 総合評価



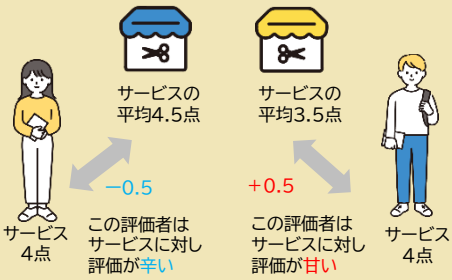

※今回は複数回投稿したユーザが分析対象

### 研究の目的

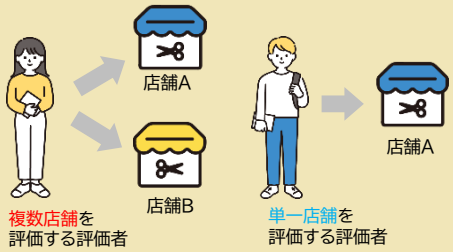
評価者の信頼性、評価者の厳しさ、評価者の傾向を把握するための特徴量エンジニアリング方法の提案、ならびにその手法による評価傾向に基づいた店舗・スタイリストのパーソナライズド推薦

## 分析のフロー（提案内容）

### Step1. 特徴量エンジニアリング



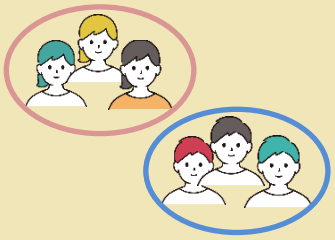

評価店舗が異なるので、同じ4点でも意味が異なるのでは？




多くの店舗を評価する評価者と同じ店舗を評価する評価者にはどのような違いがあるのか？

投稿頻度やレビューの分量の差異にどのような傾向があるか？

### Step2. 顧客クラスタリング



Step1.の特徴量に基づいたクラスタリングにより、顧客のセグメントを獲得する

### Step3. 顧客の特性を検出

XAIなどの活用

### Step4. 顧客特性に合った店舗・スタイリスト等を推薦

特徴量にはカテゴリカル変数（年代、性別）と実数の特徴量が混在しているデータになるためクラスタリングには、

**k-prototypes**を活用する

## 分析結果

### 分析の設定

- 特徴量エンジニアリングによって構築された特徴量は全て学習前に標準化している
- クラスタ数は事前に複数確認後20に決定

### 得られた知見

- 評価項目による厳しさの差は大きくはない
- 長文レビュワーの評価傾向は甘め
- 多店舗投稿に伴い投稿数は増加（自明）
  多店舗評価者は評価傾向が平均的である
- デモグラに応じて厳しさの傾向がある　などなど

### 顧客セグメントのプロトタイプ

| | ムード | サービス | テク | メニュー | 総合 | 文字数 | 店舗数 | 投稿数 | 性別 | 年代 |
|---|---|---|---|---|---|---|---|---|---|---|
| セグメント1 | -0.48 | -0.45 | -0.43 | -0.43 | -0.49 | -0.26 | -0.33 | -0.35 | F | 3 |
| セグメント2 | 0.51 | 0.52 | 0.74 | 0.4 | 0.66 | -0.64 | -0.27 | -0.21 | F | 6 |
| セグメント3 | -0.6 | -0.58 | -0.52 | -0.5 | -0.59 | -0.93 | -0.51 | -0.19 | M | 5 |
| セグメント4 | -0.36 | -0.33 | -0.33 | -0.3 | -0.35 | -0.24 | -0.28 | 2.91 | F | 8 |
| セグメント5 | 2.17 | 2.93 | 3.22 | 2.33 | 3.32 | 0.41 | 0.37 | -0.31 | F | 4 |
| セグメント6 | -0.3 | -0.41 | -0.4 | -0.4 | -0.43 | -0.29 | -0.47 | -0.22 | F | 8 |
| セグメント7 | -0.55 | -0.48 | -0.4 | -0.39 | -0.51 | -0.26 | -0.12 | -0.35 | F | 2 |
| セグメント8 | 1.37 | 1.01 | 0.11 | 0.05 | 0.49 | -0.22 | -0.05 | -0.38 | F | 4 |
| セグメント9 | -0.53 | -0.5 | -0.47 | -0.61 | -0.53 | 2.07 | -0.26 | -0.23 | F | 5 |
| セグメント10 | 0.3 | 0.12 | 0.09 | 0.28 | 0.11 | 1.54 | 0 | -0.27 | F | 3 |
| セグメント11 | 0.26 | 0.14 | 0.28 | 1.65 | 0.48 | -0.26 | -0.29 | -0.33 | F | 4 |
| セグメント12 | -0.42 | -0.45 | -0.46 | -0.51 | -0.49 | -0.17 | -0.44 | -0.26 | F | 6 |
| セグメント13 | 0.78 | 0.78 | 0.71 | 0.67 | 0.79 | 0 | 1.75 | 1.04 | F | 4 |
| セグメント14 | -0.09 | -0.22 | -0.27 | -0.27 | -0.28 | -0.27 | -0.37 | -0.27 | F | 5 |
| セグメント15 | -0.43 | -0.43 | -0.42 | -0.51 | -0.48 | -0.14 | -0.34 | -0.29 | F | 4 |
| セグメント16 | 0.04 | 0.01 | 0 | -0.03 | 0 | 0.03 | 4.13 | 3.42 | F | 4 |
| セグメント17 | -1.37 | -1.31 | -1.13 | -0.93 | -1.25 | -0.42 | -0.37 | -0.23 | F | 5 |
| セグメント18 | 1 | 1.42 | 1.61 | 1.06 | 1.63 | 1.64 | 0.22 | -0.29 | F | 6 |
| セグメント19 | -0.38 | -0.4 | -0.4 | -0.41 | -0.44 | -0.1 | 1.55 | 0.8 | F | 3 |
| セグメント20 | 1.35 | 1.56 | 1.64 | 1.46 | 1.7 | -0.47 | 0.08 | -0.29 | F | 3 |

**セグメント3**
- 30代前半男性
- 短文レビュー、店舗は一定
- 評価は全体的に厳しめ

**セグメント4**
- 40代女性
- 短文だが一部店舗に多く投稿
- 評価はやや厳しめ

**セグメント5**
- 20代後半女性
- 長文レビュー、多数店舗
- 全体的にかなり甘め

**セグメント16**
- 20代前半女性
- 多店舗に多数投稿
- 評価は平均的・ニュートラル

## 今後の展望

- BERTopicによってレビュー文から特徴量を構築
  スタイリストのPR文や店舗紹介文の分散表現をクラスタリング
- ProtNetやMMD-Criticなどの説明可能なクラスタリングを適用

**同様な評価傾向の顧客に対して高評価**な店舗・サービス・スタイリストを推薦

### 参考文献

[1]豊田ら"レビュー投稿者の甘辛度を考慮した美容院評判情報と利用者情報の可視化"第79回情報処理学会全国大会講演論文集, 521-522, (2017)
[2]岩崎ら"トピックグラフの中心性に基づいた顧客ニーズの地域差の可視化" 電学論 C, 電子・情報・システム部門誌 142 (6), (2022)
[3]平尾ら"クチコミレビューデータを用いたヘアサロンの満足度に対する特徴分析"経営情報学会　全国研究発表大会要旨集(2024)