

# 性格特性を考慮した高パーソナライズ性推薦システム

理学専攻・情報科学コース2040632 伊藤桃

## 1 はじめに

近年 AI を使ったユーザに対する推薦システムが増えてきた。主流の既存推薦システムは例えばユーザの趣味嗜好情報に基づくものであり、単に似た趣味嗜好のものを推薦する傾向がある。これはユーザにとっては単一的な推薦になってしまう。本研究では、後述するように Personality Insights [1](以下 PI) を用いたユーザ性格特性の情報提供システムを目指す。情報提供を受けた第三者は性格特性情報を用いて、スポット推薦システムやイベント推薦システムなどに応用することができ得る。本論文では、まず前提となる、スポット・イベントごとにユーザの性格特性に傾向があることを検証する。

また昨今新型コロナウイルスが、世界中の様々な業界に多大な影響を与え続けている。イベント業界においては、オンラインと現地参加いずれかの方法の参加が選択できるハイブリットイベントが新たにコロナ禍に現れた。本論文では、ハイブリットイベントに参加するユーザの参加形態別に検証することで、コロナ禍の日本のイベントに参加するユーザの性格値傾向も共に検証する。

## 2 PI が推定する性格特性

PI が推定する性格特性は性格の基本的な次元が5つであるという Big Five Model の特性項目と、Kevin Ford の Universal Needs Map に沿った Needs (欲求) 分析, Schwartz の価値概説 (Schwartz Value Survey) に沿った Values (価値観) 分析の特性項目からなる。

## 3 スポットごとの性格特性傾向の検証

今節では、pyclustering ライブラリの x-means を使い、各スポットを訪れたユーザの性格値をもとにクラスタリングする。分析に用いるデータセットは、まず Web サイト「GO TOKYO」[2] に掲載されたスポットをキーワードにツイートを検索した。そのキーワードを含めたツイートを取得期間中に行なったユーザの中で、目視でツイート内に「行った」という記述がある、あるいは実際に訪れないと知り得ない情報を含めているような日本人ユーザに絞り、そのユーザが過去にツイートしたものを先述の PI に通し、得た各特性の数値である。取得期間は 2020 年 10 月 22 日, 2020 年 12 月 20 日, 2021 年 2 月 22 日である。

結果は表 1 の通りである。

表 1 クラスタ番号とそのスポット群

クラスタ番号	スポット名
0	'神田明神', '大丸東京店', 'グランスタ', '丸の内オアゾ', '築地本願寺', '築地場外市場', '護国寺', 'ビックカメラ池袋', '不忍池', '国立科学博物館', '舎人公園', '光が丘公園', '隅田川', '江戸東京博物館', '荒川河川敷', 'ららぽーと豊洲', 'ゲートブリッジ', '豊洲ぐるり公園', '本郷公園', '葛西臨海公園', '渋谷ストリーム', '井の頭恩賜公園', '中野ブロードウェイ', '結公園', '日本科学未来館', '城南島海浜公園'
1	'乃木神社', 'サンシャイン水族館', 'アニメイト池袋', 'サンシャインシティ', 'スカイツリータウン', '渋谷ロフト', '渋谷センター街', '渋谷マークシティ', 'WOMB', '竹下通り', '表参道ヒルズ', 'アクアパーク品川', '愛宕神社', 'リキッドルーム'
2	'GINZASIX', '東急プラザ', '新丸ビル', '丸ビル', 'COREDO 室町', '日枝神社', 'コンラッド東京', '西武池袋本店', '新宿ゴールデン街', '目黒川', '六本木ヒルズ', '等々力溪谷'
3	国立映画アーカイブ
4	'秋葉原電気街', 'アニメイト秋葉原', 'ポケモンセンターメガトウキョー', 'ダイバーシティ東京', 'デックス東京ビーチ', 'アクアシティお台場'
5	'銀座三越', '永天宮', '日本橋三越', '飛鳥山公園', '吾妻橋', '京王百貨店 新宿'
6	'三菱一号館美術館', '日比谷シャンテ', '赤坂 SCT シンター', '東京芸術劇場', '日芸術劇場', '上野の森美術館', '増上寺'
7	'東京ステーションギャラリー', '根津美術館', '国立新美術館', '東京都庭園美術館'

また、それぞれのクラスタリンググループのセンターとなる性格値を図 1 のように示す。

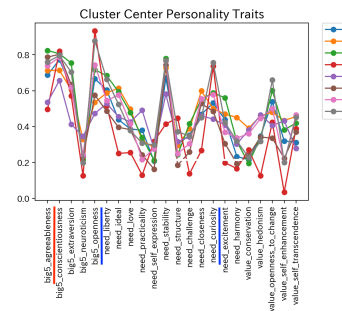


図 1 クラスタグループセンター性格値

まず、図 1 の赤線部分項目である big5\_agreeableness を見てみると、クラスタ 3,4 のグループが他グループに比べ低いことがわかる。この性格値は協調性を表し、値が低い人は人より自分の興味を優先するタイプを表す。クラスタ 3 や 4 のグループに属すスポット群を見ると、映画館、アニメイト、ポケモンセンターなど自身の趣味により訪れるスポットが多い。また、ダイバーシティ、アクアシティ、電気街のスポットも、趣味に関するイベントに参加する目的で訪れたとみれるユーザが多数見受けられ、そのようなユーザの趣味に関するツイートが値の低さにつながっているのではないかと考察する。

また、青線部分項目である need\_curiosity や big5\_openness と言った性格値がクラスタ 3,7 のグループが比較的高い。先程同様映画館や、クラスタ 7 のスポット群は美術館のスポット群であり、このようなスポットに訪れるユーザが探究心や知的的好奇心が比較的高いことは感覚的に腑に落ちる。このように PI によって求めた性格値で、クラスタリングをするとある程度感覚的に腑に落ちる形で結果が得られ、提案システムの実現ができ得ると考える。

#### 4 イベントごとの性格特性傾向の検証

まず、イベントカテゴリを「映画」「舞台」「ライブ」「美術」と設定し、各カテゴリに直近開催のイベント 10 個をマッピングし、Twitter の公式 API [3] を用いて、データを収集した。そして、前節同様の手順で各イベントのツイートをしているユーザの過去ツイートから PI 値を求め、またその PI 性格特性毎の平均値を求めた。これを後述のクラスタリングのインプットとした。各イベントカテゴリの設定は、先行研究の今井らの研究結果 [4] を参考にした。集計ユニークユーザ数は全てのカテゴリで 5 人以上である。データ取得日は、2020 年 12 月 30 日である。

先述の結果を用い、イベント毎クラスタリングを図 2 の通り実施した。横軸は性格特性、縦軸には各イベントを並べてヒートマップとして表した。色が濃いほど性格値数値が高いことを示す。また、図中の縦軸の色分けはイベントカテゴリを表しており、水色は「舞台」、紫色が「映画」、緑色が「美術」、赤色が「ライブ」カテゴリを表す。

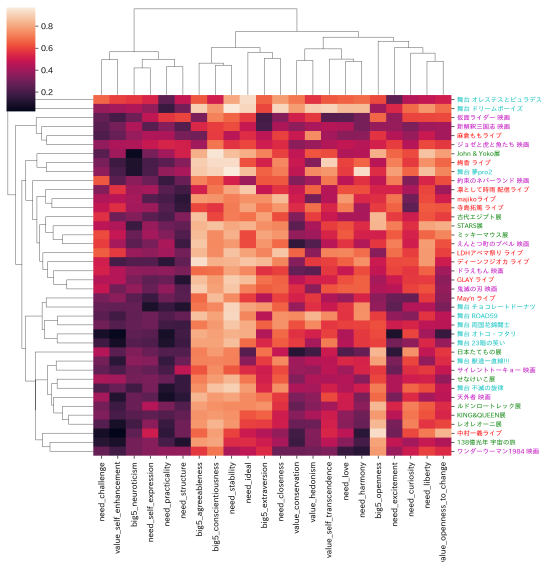


図 2 イベント別ヒートマップクラスタリング

イベントカテゴリ毎に大まかにクラスタリングされている印象を受ける。特に舞台は固まっており、舞台のイベントに参加するユーザの性格は似通っていると言える。対して、映画カテゴリはまとまりなくクラスタリングされている。映画のジャンルによって、見に行くユーザ層にばらつきがあるのではないかと考える。

#### 5 コロナ禍イベント分析

前節と同様の手順でハイブリットイベント名と参加ユーザの PI 値を取得した。次に目視で現地参加とオンライン参加ユーザに分け、各性格項目毎に箱髷図に表した。その結果 Needs (欲求) 分析の 1 つの指標である need\_love が図 3 で示す通り、比較的現地参加の方がオンライン参加より比較的高かった。この

need\_love 性格特性は外交的な性格を表す。実際に目視で現地参加とオンライン参加者のツイートを見てみると、現地参加ユーザは多数の他ユーザとやりとりを行っており、共に現地参加をしている旨のツイートもいくつか見受けられた。このようなことから、現地参加ユーザの方が need\_love 性格値が高い傾向にあることは腑に落ちる結果であると言える。図中末尾 ‘\_real’ は現地参加を ‘\_online’ はオンライン参加を表す。

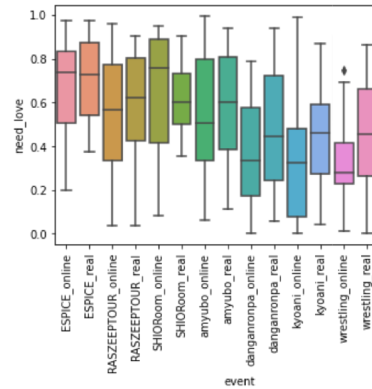


図 3 各イベント毎 ‘need\_love’ 性格項目箱髷図

また、マンホイットニーの U 検定を行うと、p 値 0.055, Hedges の g の効果量を計算すると効果量 0.20 という結果が出ており、統計的にもわずかに傾向があると言える。

#### 6 まとめと今後の課題

本研究では、ユーザの性格特性情報提供システムを実現すべく、前提となるスポット・イベントごとの性格の違いを確認すべく検証を行なった。その結果、スポットやイベントどちらもユーザ性格特性の傾向の偏りが感覚的に腑に落ちる形で確認できた。

またコロナ禍においてのスポット・イベント毎のユーザ性格値についても調査を行なった。イベントの調査についてはハイブリットイベントに注目し、オンライン参加と現地参加のユーザに分け性格特性値を分析した結果、一部の性格特性に関して両者に差があることを確認した。総じてスポット・イベント共に性格特性の傾向がいくつか見られたので、今後は具体的に提案したシステムの実現に向けて研究を進めていきたい。

#### 文 献

- [1] Personality Insights. <https://www.ibm.com/blogs/smarter-business/business/personality-Insights/>. (2021 年 12 月 29 日閲覧).
- [2] GO TOKYO. <https://www.gotokyo.org/>. (2021 年 12 月 29 日閲覧).
- [3] Twitter Search API. <https://dev.twitter.com/rest/public/search>. (2021 年 12 月 29 日閲覧).
- [4] 今井美希, 榎美紀, 工藤瑠璃子, 小口正人ほか. ソーシャルメディアを用いた訪日外国人へのパーソナライズされたイベント情報提供の最適化ルートの推薦. 第 82 回全国大会講演論文集, Vol. 2020, No. 1, pp. 391–392, 2020.