

# 卒業論文

## 特定の特徴を持つ一部の少量利用者データに対して Contextual Bandit と他利用者データを活用した観光 推薦システム

尾関 采音

2025年2月5日

岐阜大学 工学部 電気電子・情報工学科 情報コース

鈴木研究室

Title1

本論文は岐阜大学工学部に  
学士（工学）授与の要件として提出した卒業論文である。

尾関 采音

指導教員：

鈴木 優 准教授

# 特定の特徴を持つ一部の少量利用者データに対して Contextual Bandit と他利用者データを活用した観光 推薦システム\*

尾関 采音

## 内容梗概

本研究では、同じ属性の利用者が少ない利用者に対して観光地推薦システムを構築することを目的とする。

推薦システムは、利用者の嗜好に基づいて、最適な商品やコンテンツ、サービスを推薦することを目指す。利用者の嗜好を把握するためには通常、購買履歴や閲覧履歴などのデータが活用される。推薦システムを実現するための方法の一つは、協調フィルタリングであり、利用者と似た嗜好を持つ人は似たアイテムを好むという仮定のもと、利用者間や商品間の類似度を計算して新しい商品を推薦する手法である。

一方、本研究では、利用者が選択した条件を考慮する必要があるため、同じ属性の利用者があらかじめ限定される。その結果、限定された利用者の中に利用者と嗜好が類似している利用者はさらに限定される。類似利用者が少ないと推薦システムの精度が下がってしまう可能性があるというコールドスタート問題がある。

本研究では、利用者と嗜好が類似していない他の利用者の行動履歴データの影響力を抑えるために 0 以上 1 以下の値である  $\alpha$  倍の重み付けすることによってデータ不足による推薦精度の低下を防ぐことを目指す。さらに、我々は利用者が選択した条件に該当しない利用者の中にも利用者と嗜好が似ている人が含まれている可能性が高いと考えたため、利用者の嗜好を把握するために利用者とは異なり属性の利用者データも活用する。また、我々は利用者個人に特化した観光地を推薦するために Contextual Bandit と利用者からのフィードバックを用いる。

---

\*岐阜大学 工学部 電気電子・情報工学科 情報コース 卒業論文, 学籍番号: 1213033030, 2025 年 2 月 5 日.

我々は、Contextual Bandit を行う際には、事前に利用者にとどのような観光地が好みであるのか実際に選んだ結果を利用して最初にどのような特徴量を持つ観光地が利用者に適合している確率が高いのかを示す事前分布を作成することで、より精度の高い Contextual Bandit アルゴリズムを実行することが可能であると考えた。また、我々は利用者からのフィードバックを受けて確率分布を更新していくことで、より利用者個人に特化した観光地推薦が可能であると考えた。一つ目の評価実験により本提案手法は同じ属性の対象者のみのデータを用いた場合と比べて、有意差はでなかったものの平均適合率は本提案手法の方が高いことが確認できた。二つ目の評価実験により本提案手法は作成した事前分布を用いない Contextual Bandit を用いた場合と比べて、有効性が確認できた。三つ目の評価実験によりユーザベースの協調フィルタリングを用いた時と比べて、本提案手法の Contextual Bandit の初回試行の結果が、利用者によっては有効である場合もあることが確認できた。

## キーワード

Contextual Bandit, 重み付け, 情報推薦, 機械学習

# 目次

図目次	v
表目次	vi
第 1 章 はじめに	1
第 2 章 基本的事項	4
2.1 協調フィルタリング	4
2.1.1 協調フィルタリングの仕組み	4
2.1.2 協調フィルタリングの種類	6
2.1.3 協調フィルタリングの欠点	6
2.2 バンディットアルゴリズム	6
2.3 Contextual Bandit	7
2.3.1 最適なアームを求める	8
2.3.2 事後分布を求める	9
2.4 クラスタリング	11
2.4.1 $k$ -means 法	12
2.5 加重平均	13
2.6 word2vec	13
2.7 TF-IDF	14
2.8 2 標本の $t$ 検定	15
2.8.1 有意水準と $p$ 値	15
2.8.2 対応のない 2 標本の $t$ 検定	16
2.8.3 対応のある 2 標本の $t$ 検定	17
第 3 章 関連研究	18
第 4 章 提案手法	20
4.1 観光地クラスタリングのための特徴量	22
4.2 選択肢限定のためのクラスタリング	22

4.3	対象者の重み付けと平均評価値算出 . . . . .	23
4.4	作成した事前分布を利用した Contextual Bandit . . . . .	23
4.5	非同属性対象者の行動履歴データを利用者への観光地提案処理 . . .	24
<b>第 5 章</b>	<b>評価実験</b>	<b>26</b>
5.1	データセット . . . . .	26
5.2	実験手順 . . . . .	27
	5.2.1 ユーザベース協調フィルタリング . . . . .	27
5.3	結果・考察 . . . . .	28
<b>第 6 章</b>	<b>おわりに</b>	<b>31</b>
	<b>謝辞</b>	<b>33</b>
	<b>参考文献</b>	<b>34</b>
	<b>発表リスト</b>	<b>36</b>

## 図目次

4.1	手法手順 . . . . .	21
4.2	標準正規分布 . . . . .	24
4.3	利用者に事前に興味のある観光地を選んでもらって作成した事前 分布 . . . . .	24



## 表目次

2.1	利用者の各観光地に対する評価 . . . . .	5
5.1	使用データ数 . . . . .	27
5.2	各実験の適合率と $t$ 検定を行った際の $p$ 値 . . . . .	28
5.3	各実験の $t$ 検定を行った際の $p$ 値 . . . . .	28

## 第1章 はじめに

本研究の目的は、同じ属性の利用者が少ない利用者に対して観光地の推薦を行うことである。本研究で想定している利用者は、旅行において高額な出費をする利用者である。

推薦システムを実現するためには、協調フィルタリングと呼ばれる、利用者の行動データや嗜好に基づき、他の利用者との類似性を活用して個人に適したものを推薦する手法がよく使われる。[1][2]

従来のアイテム間の類似度のみを利用する手法と比べ、利用者に適した推薦が可能であるという利点が存在する一方で、利用者の行動履歴データ数が少数の場合にシステムは新しい利用者与其他の利用者の類似性が評価できないこと、行動履歴の少ない新規利用者に適切な推薦ができないことが欠点として挙げられる。

観光に着目すると、ある属性の利用者に限定して観光地を提案しようとする、同じ属性の利用者だけでなく同じ嗜好を持った人、提案できる観光地も数が限定されてしまう。以上の理由より、我々は本研究の目的のために協調フィルタリングよりも高い適合率の推薦システムを構築することができないかと考えた。また、本研究では類似利用者の少なさが問題であるため、行動履歴データの量を補う必要があると考えた。

我々は、本研究の問題点である類似利用者の少なさを解決するために、利用者嗜好が異なる利用者の行動履歴データを重み付けすることによる活用でデータを補填し、利用者の嗜好を少量データをもつ特定の条件に該当する利用者に対しても行動履歴データ量が充分ある場合と比べて十分な精度を持つ観光推薦システムを構築することを目指す。

提案手法では、主に二つの処理を行う。一つ目は、同じ属性の利用者以外の利用者のデータを利用することである。同じ属性の利用者以外の利用者の行動履歴データを利用することによって新規利用者の嗜好に近い嗜好を持つ人が含まれる可能性が高くなるため、推薦システム利用者の嗜好をより正確に把握し、それに基づいた適切な推薦が可能であると考えられる。条件に該当する利用者以外の利用者の行動履歴データは、利用者嗜好が違う他の利用者に0以上1以下の値である $\alpha$ 倍の重み付けをして加重平均を求め、算出した結果を観光地の新しい平均評価値とするこ

とで利用する。これにより、嗜好が異なる人の行動履歴データの影響力を抑えることが可能である。結果、利用者に適合する観光地を推薦しやすくなると考えた。

二つ目は、Thompson Sampling[3] を用いた Contextual Bandit による観光地推薦と利用者からのフィードバックを行うことである。我々は利用者から提案された観光地に対する興味の有無についてのフィードバックをもらい、Thompson Sampling を用いた Contextual Bandit による探索と活用を繰り返すことで、より個人に特化した精度の高い推薦システムを構築することが可能であると考えた。また、我々は、Contextual Bandit を実行する前に、利用者実際に興味があるカテゴリを選んでもらい、その結果をもとに事前分布を作成することによって、精度をさらに向上させることが可能であると考えた。

我々は、提案手法の有効性を確かめるために評価実験を行った。実験では、被験者に利用者が希望している条件を利用者設定として提示し、被験者に対して観光地推薦を行う利用者実験を行った。比較実験として三つの実験を行った。一つ目は、条件に該当する他の利用者の行動履歴データのみを扱った Contextual Bandit を比較手法とした実験である。実験の目的は、条件に該当しない他の利用者の行動履歴データを扱うことの有効性を確かめることである。二つ目は、事前分布を作成せずに標準正規分布を用いた Contextual Bandit を比較手法として行った実験である。実験の目的は、作成した事前分布を用いたことによる精度向上を確かめることである。三つ目は、ユーザベース協調フィルタリングを用いた観光推薦システムを比較手法とした実験である。実験の目的は、ユーザベース協調フィルタリングを用いた観光地推薦システムと比較して、提案手法を用いたシステムの方が利用者が興味を持つ観光地をより多く提案できるかどうかを確認することである。

実験の結果は、すべての実験において有意差は確認できなかった。しかし、条件に該当しない他の利用者の行動履歴データを扱うことの有効性と、作成した事前分布を用いたことによる精度向上、協調フィルタリングより本提案手法が適合率が高い場合があることを確かめることができた。

以下、本手法を利用する人を利用者、利用者以外の人を対象者、利用者が希望している条件に該当する観光地に訪れている対象者を同属性対象者、同属性対象者以外を対象者を非同属性対象者、属性が同じかどうかにかかわらず利用者と嗜好が類似している対象者を類似対象者、利用者と嗜好が類似していない対象者を非類似対

象者と呼ぶ.

本研究の貢献は以下の通りである.

- 非同属性対象者の行動履歴データの活用により十分な量のデータである場合と同等の精度を出すことが一部の利用者において可能であることが分かった.
- 作成した事前分布を活用した Contextual Bandit と利用者からのフィードバックの適用が個人に特化した推薦システムの提案に有効である場合もあることが分かった.

本論文の構成は以下の通りである. 2 章では関連研究について述べる 3 章では, 提案手法の Contextual Bandit と行動履歴データの重み付けを用いた観光推薦システムについて述べる. 4 章では比較実験 3 種類と提案手法による利用者実験の合計 4 種類の実験について述べる. 最後に 5 章では本論文のまとめと今後の課題について述べる.

## 第 2 章 基本的事項

### 2.1 協調フィルタリング

推薦システムに用いられるアルゴリズムは、代表的なもので協調フィルタリングとコンテンツベースフィルタリング、協調フィルタリングとコンテンツベースフィルタリングを組み合わせたハイブリットがある。協調フィルタリングは、利用者間の行動データやアイテムの評価データを利用して、個々の利用者に適したアイテムを推薦するシステムである。コンテンツベースフィルタリングは、事前に観光地を属性に応じてグループ化しておき、利用者が検索したり訪れた観光地と似た観光地を特徴が似ている順に表示して推薦するシステムである。コンテンツベースフィルタリングを利用する場合、あらかじめ用意された観光地の情報と利用者の嗜好を比較し、類似度の高い観光地をおすすめする。そのため、必然的におすすめされる観光地は、利用者がすでに知っている観光地であることが多い。一方、協調フィルタリングでは、利用者と似た嗜好を持つ他の人々の行動履歴を利用するため、利用者自身の情報が少ない場合でも比較的興味に合った観光地をおすすめすることが可能である。また、類似利用者の行動履歴に含まれている観光地で、利用者自身はまだ訪れていないものをおすすめ候補とすることから、利用者が初めて知る観光地を提案できる可能性がある。

#### 2.1.1 協調フィルタリングの仕組み

本節では本研究の協調フィルタリングの基本的な仕組みについて説明する。

- 1 被験者と対象者が訪れたら 1、訪れていないならば 0 の評価をつけたものの表を作成する。
- 2 被験者と対象者間の類似スコアをコサイン類似度で算出する。

作成する表は表 2 のようなものである。ただし、 $r_{u,i}$  は利用者  $u$  の観光地  $i$  に対する訪れたかどうかの判定であり、訪れたことがあれば 1。訪れたことがなければ 0 であるとする。

表 2.1 利用者の各観光地に対する評価

観光地名 利用者名	観光地 1	観光地 2	...	観光地 $m$
利用者 1	$r_{1,1}$	$r_{1,2}$	...	$r_{1,m}$
利用者 2	$r_{2,1}$	$r_{2,2}$	...	$r_{2,m}$
...	...	...	...	...
利用者 $n$	$r_{n,1}$	$r_{n,2}$	...	$r_{n,m}$

表 2 を元にした評価行列  $R$  は (2.1.1) 式で表すことができる.

$$R = \begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,m} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n,1} & r_{n,2} & \cdots & r_{n,m} \end{pmatrix} \quad (2.1.1)$$

類似度の測定方法は主に以下の 2 通りである.

- 1 コサイン類似度
- 2 ピアソン相関係数

コサイン類似度は以下の式で表せる. ただし,  $I_{uv}$  は利用者  $u$  と  $v$  が共通して評価した観光地の集合である.

$$\cos(u, v) = \frac{\sum_{i \in I_{uv}} r_{u,i} r_{v,i}}{\sqrt{\sum_{i \in I_{uv}} r_{u,i}^2} \sqrt{\sum_{i \in I_{uv}} r_{v,i}^2}}$$

ピアソン相関係数は以下の式で表せる.

$$\text{sim}(u, v) = \frac{\sum (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum (r_{ui} - \bar{r}_u)^2} \sqrt{\sum (r_{vi} - \bar{r}_v)^2}}$$

本研究では 0 から 1 の範囲に類似度の値が収まり, 直感的に類似性を解釈しやすいコサイン類似度を利用する.

### 2.1.2 協調フィルタリングの種類

協調フィルタリングにはメモリベースとモデルベースがある。メモリベースにはユーザベース協調フィルタリングとアイテムベース協調フィルタリングがある。ユーザベース協調フィルタリングは類似する嗜好を持つ他の利用者が訪れた観光地を推薦する。具体的には、2.1.1節で求めた類似度が最も大きい  $n$  人の利用者を選択し、類似度を重みとして評価の加重平均を求め、利用者に観光地を推薦する。加重平均は以下の式の通りである。

$$r_{ui}^- = \frac{\sum_{v=1}^n w_v r_{vi}}{\sum_{v=1}^n w_v}$$

アイテムベース協調フィルタリングは、利用者が訪れた観光地と特徴量が類似する他の観光地を推薦する。本研究では、提案手法の有効性を確かめるために比較実験でユーザベース協調フィルタリングを用いる。

### 2.1.3 協調フィルタリングの欠点

協調フィルタリングの欠点はコールドスタート問題と呼ばれる、新しい利用者や誰も行ったことがない観光地がシステムに入力されたときに推薦精度の低下が生じる問題がある。利用者が観光地に訪れたかがどうかについて履歴がないため、システムは新しい利用者の類似性や既存の利用者との関連性を評価できない。また、行動履歴の量がある程度存在する訪問者でも、嗜好が特殊で類似の利用者を発見することが困難な場合、おすすめ観光地を見つけるデータがないのと同じ状況であり、協調フィルタリングでの推薦が困難である。

## 2.2 バンディットアルゴリズム

バンディットアルゴリズムは、限られた情報の中で最適な選択肢を見つけ出すためのアルゴリズムである。[4] バンディットアルゴリズムの目的は、試行錯誤を繰り返しながら得られる成果 (報酬) を最大化することである。

例として、推薦システムと広告を挙げる。推薦システムでは、ウェブにおける

サービスサイトにおいて、過去の購買履歴に基づいて各利用者の訪問時に全商品の中からいくつかの商品を推薦し、推薦した商品のうち実際に購入される商品数または購入による利益の最大化を目指す。

バンディットアルゴリズムは、得られる成果を最大化するために探索と活用を繰り返している。探索とはどの選択肢を選ぶとどのような成果が得られるかという情報を獲得することで、活用とは獲得した情報を使いより良い選択肢を選択することである。推薦システムで例えると、探索はまだ利用者に推薦したことがない類の商品を推薦し、購入されるかどうかを確かめることで、活用は以前利用者が購入した商品と類似する商品を推薦することである。

選択肢が複数ある場合を多腕バンディット問題と呼ぶ。多腕バンディット問題は、一つひとつの選択肢を腕 (arm) として考え、未知の報酬分布を持つ  $K$  本の腕を逐次的に選択し、その累計報酬を最大化することが目的である。

## 2.3 Contextual Bandit

Contextual Bandit とは、事前に選択肢を選んだことで得られる報酬を仮定した確率分布のパラメータを利用者の特徴といった Context に応じて変化させるアルゴリズムである。

推薦システムにおけるバンディットアルゴリズムの目的は、推薦した商品のうち実際に購入される商品数または購入による利益の最大化であるが、実際は利用者の層によって、効果の高い推薦システムは変わってくる。利用者が購入した商品情報 (Context) を考慮して、それに応じた報酬分布を仮定した Contextual Bandit を利用する。

本研究では Contextual Bandit アルゴリズムを用いてどのようなコンテキストを持つ観光地が利用者に気に入ってもらえる確率が高いのか予測している。本研究では、報酬が 2 値である Thompson Sampling を用いた Contextual Bandit アルゴリズム [3] を用いる。

Thompson Sampling とは、どの選択肢が一番高い確率で報酬を得られるかを確率分布から求めるものである。手順は以下の通りである。

1. 各選択肢に対して確率分布を表すパラメータを初期化する。一般的にはベ-



- タ分布などの確率分布が使用される。
2. 各選択肢に対して現在の確率分布からサンプリングを行い，それに基づいて選択肢を選ぶ．これにより，より報酬の高い選択肢が選ばれやすくなるが，不確実性を考慮して探索も行われる．
  3. 選ばれた選択肢から得られた報酬を受け取る．
  4. 得られた報酬を用いて各選択肢の確率分布のパラメータを更新する．

報酬が2値である Thompson Sampling を用いた Contextual Bandit アルゴリズムの処理手順について [4] を参考に説明する．ただし，各アームを  $a_1, a_2, a_3, \dots, a_n$ ， $i$  を  $1, 2, \dots, n$  とする． $t$  回目に  $i$  番目のアーム  $a_i$  を提示したとすると，利用者やアーム  $a_i$  の  $t$  回目のコンテキストを  $x_{t,a_i}$  とする．

1. 利用者にアーム  $a_i$  を提示する．
2. 利用者の特徴や  $i$  のアーム  $a_i$  のコンテキスト  $x_{t,a_i}$  と利用者による興味の有無， $t$  回目までの利用者からのフィードバックを用いて，利用者が興味を持つと考えられるアームを推定するために事後分布を算出する．
3. 事後分布から反復を繰り返すニュートン法により MAP 推定量  $\hat{\theta}_t^{MAP}$  を求める．
4. 求めた MAP 推定量  $\hat{\theta}_t^{MAP}$  を用いて事後分布を近似し，事後分布からパラメータをランダムにサンプリングする．
5. 4 のパラメータを用いて最適なアームを求める．

### 2.3.1 最適なアームを求める

$t$  回目のフィードバックを受けた時のパラメータを  $\theta_t$ ，各アームの報酬を  $r_{t,a_i}$  とすると，次に利用者が行きたいか否かフィードバックされる各観光地の報酬の確率はそれぞれ以下の式になる．報酬とは，利用者からフィードバックされる興味の有無のことである．報酬が二値であるためロジスティック回帰を使用する．以降，利用者の興味ありを 1，興味なしを 0 で表現する．

$$P(r_{t,a_i}|\theta_t) = \begin{cases} \frac{e^{\theta_t^T x_{t,a_i}}}{1+e^{\theta_t^T x_{t,a_i}}} & (r_{t,a_i} = 1) \\ \frac{1}{1+e^{\theta_t^T x_{t,a_i}}} & (r_{t,a_i} = 0) \end{cases}$$

利用者が行きたいと思うアーム (観光地) を求めたいため, 上記より  $t$  回目のフィードバック後の最適なアームは次のように求められる.

$$a(t) = \arg \max_{a_i} \frac{e^{\theta_t^T x_{t,a_i}}}{1+e^{\theta_t^T x_{t,a_i}}} = \arg \max_{a_i} \theta_t^T x_{t,a_i} \quad (2.3.1)$$

### 2.3.2 事後分布を求める

2.3.1 の最適なアームを導くために線形パラメータ  $\theta_t$  の事後分布を導出する必要がある. ロジスティック回帰モデルでは, 事後分布を解析的に求めることはできない. よって以下ではラプラス近似で線形パラメータの事後分布を変量正規分布に近似した上でサンプリングするという手順を踏む.  $t$  回目までの事後分布を求める式は以下である.

$$p(\theta_t|\{r_{s,a_i}\}_{s=1}^t) = \frac{p(\theta_t) \prod_{s=1}^t p(r_{s,a_i}|\theta_s)}{\prod_{s=1}^t p(r_{s,a_i})}$$

パラメータ  $\theta_t$  とは, 求めた事後分布からランダムにサンプリングした値である.

実際にパラメータ  $\theta_t$  を求めるために, 先ほど述べた報酬の確率を用いて事後分布の事後確率最大化推定 (MAP 推定) を行う. ただし,  $\sigma$  はハイパーパラメータであり, パラメータ  $\theta_t$  の事前分布は事前に利用者から受け取ったフィードバックを用いて作成したものとする.

$$\begin{aligned} p(\theta_t|\{r_{s,a_i}\}_{s=1}^t) &= \frac{p(\theta_t) \prod_{s=1}^t p(r_{s,a_i}|\theta_s)}{\prod_{s=1}^t p(r_{s,a_i})} \\ &\propto p(\theta_t) \prod_{s=1}^t p(r_{s,a_i}|\theta_s) \\ &\propto p(\theta_t) \prod_{s=1}^t \left( \frac{e^{\theta_s^T x_{s,a_i}}}{1+e^{\theta_s^T x_{s,a_i}}} \right)^{r_{s,a_i}} \left( \frac{1}{1+e^{\theta_s^T x_{s,a_i}}} \right)^{1-r_{s,a_i}} \end{aligned}$$

$$\propto \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\theta_t^T \theta_t}{2\sigma^2}} \prod_{s=1}^t \frac{1}{1 + e^{\theta_s^T x_{s,a_i}}} \prod_{s:r_{s,a_i}=1} e^{\theta_s^T x_{s,a_i}}$$

$\theta_t$  に関する項に注目した負の対数事後分布は (2.3.2) になる.

$$\begin{aligned} -\log p(\theta_t | \{r_{s,a_i}\}_{s=1}^t) &= \frac{\theta_t^T \theta_t}{2\sigma^2} + \sum_{s=1}^t \log(1 + e^{\theta_s^T x_{s,a_i}}) \\ &\quad - \sum_{s:r_{s,a_i}=1} \theta_s^T x_{s,a_i} + const \quad (const: \text{定数項}) \quad (2.3.2) \end{aligned}$$

以下で MAP 推定量  $\hat{\theta}_t^{MAP}$  を求める方法について述べる.

(2.3.2) 式の負の対数事後分布の勾配ベクトルとヘッセ行列は以下のように計算できる. ただし,  $I_d$  は単位行列とする.

$$\begin{aligned} G(\theta_t) &= \frac{\theta_t}{\sigma^2} + \sum_{s=1}^t \frac{e^{\theta_s^T x_{s,a_i}} x_{s,a_i}}{1 + e^{\theta_s^T x_{s,a_i}}} - \sum_{s:r_{s,a_i}=1} x_{s,a_i} \\ H(\theta_t) &= \frac{I_d}{\sigma^2} + \sum_{s=1}^t \frac{e^{\theta_s^T x_{s,a_i}} x_{s,a_i} x_{s,a_i}^T}{(1 + e^{\theta_s^T x_{s,a_i}})^2} \end{aligned}$$

前述の通り,  $G(\theta_t) = 0$  を満たす MAP 推定量解析的に求めることができない. そこで, 反復を繰り返すニュートン法により  $\hat{\theta}_t^{MAP}$  を数値的に求める. 反復を繰り返すニュートン法は以下の式で求める.

$$\hat{\theta}_t(j+1) \leftarrow \hat{\theta}_t(j) - \frac{G(\hat{\theta}_t(j))}{H(\hat{\theta}_t(j))}$$

このニュートン法により求めた MAP 推定解の周りで先ほどの負の対数尤度を 2 次近似する. ここで  $G(\hat{\theta}_t^{MAP}) \approx 0$  とすると, (2.3.3) 式になる.

$$\begin{aligned} -\log p(\theta_t | \{r_{s,a_i}\}_{s=1}^t) &\approx \frac{1}{2} (\theta_t - \hat{\theta}_t^{MAP})^T H(\hat{\theta}_t^{MAP}) (\theta_t - \hat{\theta}_t^{MAP}) + \\ &\quad \frac{(\hat{\theta}_t^{MAP})^T \hat{\theta}_t^{MAP}}{2\sigma^2} + \sum_{s=1}^t \log(1 + e^{(\hat{\theta}_t^{MAP})^T x_{s,a_i}}) - \\ &\quad \sum_{s:r_{s,a_i}=1} (\hat{\theta}_t^{MAP})^T x_{s,a_i} + const \quad (2.3.3) \end{aligned}$$

ここで、多変量ガウス分布の確率密度関数に負の対数をとると (2.3.4) 式になる。

$$-\log \frac{1}{(\sqrt{2\pi})^n \sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) = \frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) + const \quad (2.3.4)$$

(2.3.3) 式の一項目と (2.3.4) 式を比べるとほぼ同じ形をしている。よって、求めたい事後分布を多変量ガウス分布に近似する。解析的に求められない事後分布を、ニュートン法などで求めた MAP 推定量を平均に、ヘッセ行列の逆行列を分散共分散行列に持つ多変量正規分布  $MultiNormal(\hat{\theta}_t^{MAP}, (H(\hat{\theta}_t^{MAP}))^{-1})$  で近似するのがラプラス近似である。近似した事後分布から線形パラメータをサンプリングすることで、ロジスティック回帰モデルを用いた時の Thompson Sampling を構成することができる。

## 2.4 クラスタリング

クラスタリングとは、データ間の類似度に基づいてデータをグループ分けする手法である。クラスタリングには、階層的クラスタリングと非階層的クラスタリングが存在する。階層的クラスタリングとは、データ間の類似度が近いものからまとめていく手法のことである。階層的クラスタリングは、それぞれのデータ間における、顧客データにおける年齢や購買履歴などといったデータ特徴量の距離に基づいてクラスタを形成する。非階層的クラスタリングとは、データをあらかじめ指定したクラスタ数に基づいてグループ化し、それぞれのデータを、顧客データにおける年齢や購買履歴などといったデータの特徴量を基に適切なクラスタに割り当てる手法である。

非階層的クラスタリングの代表的なアルゴリズムとして  $k$ -means 法がある。非階層的クラスタリングの利点は、階層的クラスタリングと比べて計算量が少なく、大規模データセットに適用しやすい点である。また、クラスタの数を事前に設定することで、問題の性質や目的に応じたクラスタリングを実現できる。

### 2.4.1 $k$ -means 法

$k$ -means 法の手順は以下の通りである。

1. ランダムにクラスタの中心  $c_1, c_2, \dots, c_k$  を選択する。ただし、 $k$  はクラスタ数である。
2. 各データの特徴量  $x_j$  に対して、ユークリッド距離の 2 乗の大きさが最も小さいクラスタの中心  $c_i$  を割り当てる。
  - ユークリッド距離  
各データの特徴量  $x_j$  とクラスタの中心  $c_i$  のユークリッド距離の 2 乗は以下の式で求められる。

$$d(c_i, x_j) = (x_j - c_i)^2$$

3. クラスタの中心を更新する。  
更新後のクラスタの中心を  $c'_i$ 、クラスタの要素の数を  $S$  とすると更新されたクラスタの中心の位置は以下の式で表せる。

$$c'_i = \frac{1}{S} \sum x_j$$

4. 収束するまで 2, 3 を繰り返す。  
クラスタ中心の更新による位置の変化がある閾値以下になるか、またはアルゴリズムが指定された最大繰り返し回数に達した場合に収束とみなす。具体的には、すべての  $i$  に対して新旧のクラスタ中心間  $c_i, c'_i$  の距離が閾値  $\epsilon$  以下になった場合が収束である。収束の条件は以下の式の通りである。

$$(c'_i - c_i)^2 < \epsilon$$

$k$ -means 法の主な利点は、アルゴリズムが単純で計算コストが比較的安く、大規模なデータセットにも適用可能な点である。

## 2.5 加重平均

加重平均とは、各データの特徴量に異なる重みを設定し重みを加味した上で平均を求める手法である。加重平均は、すべてのデータを同等に扱う単純平均とは異なり、特定のデータに重点を置いた分析を可能にする。

具体的には、各データの特徴量を  $x_1, x_2, \dots, x_n$  に対する重みを  $w_1, w_2, \dots, w_n$  とすると、加重平均  $\bar{x}_w$  は以下の式で表される。

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

## 2.6 word2vec

word2vec[5][6] とは、文章に含まれる単語を分散表現ベクトルに変換し、単語の意味を把握するという自然言語処理の手法である。単語をベクトルで表現することで、単語間の距離や内積が定義され単語間の近さを定量化できるという利点がある。また、数値によって表現されていることから、機械学習モデルへの入力として用いることができる。

word2vec で得たベクトルには、単語の意味や文法が反映される。得られたベクトルは、単語の分散表現と呼ばれ各単語がベクトル空間上の点として分散している状態である。

word2vec のモデルには、CBOW(Continuous Bag of Word) と Skip-Gram の 2 種類のニューラルネットワークがある。CBOW は周囲の複数の単語から対象の単語を予測し、Skip-Gram は対象となる単語から周囲の複数の単語を予測する。CBOW と Skip-Gram のどちらも、コーパスを用いてニューラルネットワークを訓練し、パラメータから分散表現を獲得する。CBOW では予測する単語が 1 つであるのに対して、Skip-Gram では複数の周辺単語を予測し、学習時に複数単語の予測誤差を計算する必要があるため、計算時間がより長くなる。しかし、多くの場合、Skip-Gram で得た分散表現の方が高い精度になると知られている。

word2vec の仕組み [7] は、以下の手順の通りである。

1. 入力層で単語を one-hot ベクトル [8] として表したものを入力する。

one-hot ベクトルとは、特定の要素が 1 で、それ以外の要素がすべて 0 であるベクトルのことである。例として、"塩"を one-hot ベクトルで表すと以下の通りである。

$$x_{\text{塩}} = [1, 0, 0, 0, 0]^T.$$

2. 隠れ層で以下の式を用いて単語を埋め込みベクトルという意味や文脈を反映した数値ベクトルに変換する。ただし、 $x$  は次元数  $V$  である入力単語のワンホットベクトル、 $W$  は  $V \times N$  の重み行列、 $h$  は、 $N$  次元の隠れ層の出力ベクトルである。

$$h = W^T x,$$

ここで、重み  $W$  の各行は入力語に対する単語ベクトルである。本研究では、入力語に対する単語ベクトルを用いて単語間の類似度を求める。

3. 出力層で、以下の式の通り隠れ層の出力  $h$  に  $N \times V$  の重み行列  $W'$  を掛けて線形変換後の値  $z$  を求める。

$$z = W'h$$

また、出力層の重み行列  $W'$  の各列は、出力単語のベクトルである。

4. 線形変換後の値を softmax 関数を用いて、出力単語が  $w_i$  である確率  $P(w = w_i | h)$  を、以下の式で求める。

$$P(w_i | h) = \frac{\exp(z_i)}{\sum_{j=1}^V \exp(z_j)}.$$

## 2.7 TF-IDF

TF-IDF(Term Frequency - Inverse Document Frequency) は、文書集合内の単語の重要度を測定する指標の一つである。具体的には、文書内である単語がどれくらい多い頻度で出現するかを表す TF 値と、全文書中である単語を含む文書がどれくらい少ない頻度で存在するかを表す IDF 値を掛け合わせた値でのことある。TF-IDF 値は次の式で表すことができる

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \cdot \text{IDF}(t)$$

ここで、各要素は以下のように定義される。

- TF (Term Frequency)

$$\text{TF}(t, d) = \frac{f(t, d)}{\sum_{t' \in d} f(t', d)}$$

- $f(t, d)$ : 文書  $d$  内で単語  $t$  が出現する回数
- $\sum_{t' \in d} f(t', d)$ : 文書  $d$  内の全単語の出現回数の総和

- IDF (Inverse Document Frequency)

$$\text{IDF}(t) = \log \frac{N}{1 + \text{df}(t)}$$

- $N$ : 総文書数
- $\text{df}(t)$ : 単語  $t$  を含む文書の数

- TF-IDF の統合式

$$\text{TF-IDF}(t, d) = \frac{f(t, d)}{\sum_{t' \in d} f(t', d)} \cdot \log \frac{N}{1 + \text{df}(t)}$$

## 2.8 2 標本の $t$ 検定

本研究では、提案手法と比較手法の推薦精度に有意差があるかどうか確認するために 2 標本の  $t$  検定を用いた。本節では、対応のない 2 標本の  $t$  検定と対応のある 2 標本の  $t$  検定の手順について述べる。

### 2.8.1 有意水準と $p$ 値

統計的仮説検定では、帰無仮説を棄却するかどうかの判断基準として有意水準と  $p$  値が用いられる。



- 有意水準 ( $\alpha$ ) : 帰無仮説を棄却する基準である. 一般的には,  $\alpha = 0.05$  や  $\alpha = 0.01$  が用いられる.
- $p$  値: 観測されたデータに基づく検定統計量が, 帰無仮説の下で得られる確率を表す.  $p$  値が有意水準以下の場合, 帰無仮説を棄却する.

## 2.8.2 対応のない 2 標本の $t$ 検定

対応のない 2 標本の  $t$  検定は, 2 つの独立した群の平均値を比較する際に使用される. 手順は以下の通りである.

### 1. 仮説の設定

帰無仮説と対立仮説を設定する.

帰無仮説  $H_0$  : 二つの標本の平均に有意な差はない

対立仮説  $H_1$  : 二つの標本の平均に有意な差がある.

### 2. 有意水準の設定

本研究においては, 有意水準  $\alpha = 0.05$  とする.

### 3. 検定統計量の計算

検定統計量  $t$  は, 二つの標本の母平均をそれぞれ  $\mu_1, \mu_2$ ,  $\bar{x}_1, \bar{x}_2$  を群 1 と群 2 の標本平均,  $n_1, n_2$  を群 1 と群 2 の標本サイズとすると下の式で表すことができる.

$$\begin{aligned} t &= \frac{(\bar{x}_1 - \mu_1) - (\bar{x}_2 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ &= \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \end{aligned}$$

ただし, 帰無仮説である「母平均が等しい」という仮説が正しいとすると,  $\mu_1 - \mu_2 = 0$  となる. したがって, 検定統計量  $t$  は以下の式の通りになる.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

また、上の式の  $s^2$  は二つの標本の不偏分散を  $s_1^2$ ,  $s_2^2$  とすると以下の式の通りである.

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

#### 4. 有意水準と $p$ 値の比較

検定統計量  $t$  と  $t$  分布を基に  $p$  値を求め、有意水準  $\alpha$  より  $p$  値が下回った時に帰無仮説が棄却される. 帰無仮説が棄却されることによって、対立仮説が採択され、二つの標本の平均に有意な差があると言える.

### 2.8.3 対応のある 2 標本の $t$ 検定

対応のある 2 標本の  $t$  検定は、同一データにおける 2 つの条件間の平均値を比較する際に使用される. 手順は以下の通りである.

#### 1. 仮説の設定

帰無仮説と対立仮説を設定する.

帰無仮説  $H_0$  : 二つの標本の平均に有意な差はない

対立仮説  $H_1$  : 二つの標本の平均に有意な差がある

#### 2. 有意水準の設定

有意水準は対応のない 2 標本の  $t$  検定と同じで、 $\alpha = 0.05$  とする.

#### 3. 検定統計量の計算

二群間の差の平均を  $\bar{d}$ ,  $s^2$  を不偏分散,  $n$  を自由度とすると検定統計量  $T$  を以下の式で計算する.

$$T = \frac{\bar{d}}{\sqrt{\frac{s^2}{n}}}$$

#### 4. 有意水準と $p$ 値の比較

検定統計量  $T$  を基に  $p$  値を求め、有意水準  $\alpha$  と比較する. 判断基準は、対応のない検定と同様である. ただし、片側検定の  $p$  値は  $t$  分布における検定統計量の上側確率である.

### 第 3 章 関連研究

推薦システムにおいて、どのような手法が精度向上に有効であるのかについての研究はすでになんども行われている。

安江ら [9] は賃貸物件検索において、適合性フィードバックに Contextual Bandit アルゴリズムを適用することで適切な条件を入力できない場合でも利用者が少ない労力で賃貸物件検索を行うことができる Web アプリを作成する研究を行った。バンディットアルゴリズムの各アームを各賃貸物件、報酬を利用者が行う適合・不適合の判定と置き換えていた。また、事前分布は標準正規分布を用いて利用者からのフィードバックに応じて確率分布を更新した。また、著者が独自に設定した検索クエリと物件選択条件によるシミュレーション実験と、被験者に利用者設定を提示し、状況に応じた物件検索を行う利用者実験を行った。Rocchio の式を適用したシステムと比較し、Contextual Bandit を適用したシステムの方がフィードバック回数が少なくなる実験条件がいくつかあることを確認したが、物件選択条件による該当物件数が多いと考えられる場合においては Rocchio の式を適用したシステムでのフィードバック回数の方が少ない実験条件もいくつかあることも確認した。

本研究でも、適合性フィードバックに Contextual Bandit アルゴリズムを適用した。ただし、各アームを観光地、報酬を利用者からの行きたい、行きたくないの判定とした。また、この研究で用いられた事前分布は正規分布  $N(0, \sigma^2)$  であるが、本研究で用いた事前分布はあらかじめ利用者から受け取った観光地の選択結果を用いて作成したものであるため、一回目の提案から高精度な提案が可能であると考えた。

村上ら [10] は、CAM 法という全商品に対する商品属性ベクトルと一部の商品に付与された利用者の選考情報として利用者の選考情報選考情報を入力として、利用者の嗜好を反映した全商品がマップされたベクトル空間を生成することが可能な属性生成手法を用いて個人の嗜好や主観を反映したモデルを用いた推薦システムを構築し、利用者によるアンケート評価を用いてシステムの有効性を確認している。

土方 [11] によるとユーザプロファイリングには利用者から直接興味に関する情報を入力してもらう明示的手法と、利用者の Web 閲覧時の挙動閲覧時の挙動から利用者の興味に関する情報を取得する暗黙的手法に分類できる。

本研究は、利用者どの観光地を選んだかという行動履歴を用いて利用者の嗜好を

推定する暗黙的手法を用いる。

三宅ら [12] は、商品推薦手法の有効性は様々な文脈に左右されるため、実環境での継続的な評価なしに最適な推薦手法を選択することは困難であると考え、推薦手法の選択を自動的かつ継続的に最適化する推薦システムを提案した。三宅らが提案システムでは、最善な推薦手法の選択を多腕バンディット問題とみなして解いた。また、商品カテゴリや時間帯といった推薦手法の選択に影響を及ぼす利用者や商品、環境の特性を文脈として考慮することで推薦手法の選択を最適化した。評価実験の結果、EC サイトにおいて時間の経過並びに商品特性の差異により有効な推薦手法が異なること、そのため時間による推薦手法の有効性の変化並びに閲覧中の商品カテゴリを文脈とみなした最適化手法が、文脈を考慮しない場合と比較して累積クリック数を約 2% 増加させる効果があることを確認した。

本研究でも文脈を考慮したバンディットである Contextual Bandit を用いて観光地推薦システムを構築するが、推薦手法の選択のためではなく、利用者にとって最適な観光地を推薦するために Contextual Bandit を用いる。また、レビューの TF-IDF ベクトルとカテゴリの単語ベクトルと対象者の観光地への評価値を文脈 (コンテキスト) として活用する。

## 第 4 章 提案手法

我々は、非同属性対象者の行動履歴データを活用するために、類似嗜好対象者ならば同じ属性でなくても利用者が訪れたいと思う場所の評価値やカテゴリといった特徴が類似するのではないかという仮説を立てた。提案手法において必要な行動履歴データは、対象者名と対象者が訪れた観光地名、対象者の観光地に対するレビューと総合評価値、観光地のカテゴリ、観光地の詳細が記載された URL である。提案手法の流れは図 4.1 で示す。

提案手法の手順は事前分布の作成手順と利用者への観光地推薦手順の二つに大きく分けることができる。事前分布作成手順は以下の 1. から 5. で、利用者への観光地推薦手順は 6. から 8. である。具体的な手順は以下の通りである。

提案手法の手順は以下の通りである。

### 1. 利用者が興味のあるカテゴリを選択

高級ホテル、温泉、ビーチのようなカテゴリから興味のあるカテゴリを利用者が複数選択する。

### 2. 観光地のクラスタリング

1. で選ばれたカテゴリに該当するデータのうち、カテゴリ、観光地に訪れた対象者のレビューを特徴量としてクラスタリングを行う。

### 3. 各クラスタ代表観光地から行きたい観光地を選択

2. の各クラスタにおいて観光地における対象者の平均評価値が一番大きい観光地を利用者に提示し、行きたい観光地番号を入力する。

### 4. 全データの評価値に重み付け

3. で選んだ観光地に訪れていない対象者の全評価値に重み  $\alpha$  を  $0 \leq \alpha < 1$  の範囲で付ける。その他の対象者には重み 1 をつける。その後、全ての観光地において総合評価値の加重平均を求め、新たな平均評価値とする。これにより、利用者と非類似嗜好対象者の行動履歴データを無視することなく影響を抑えることができる。

### 5. 事前分布作成

3. の結果を用いてレビューの TF-IDF 値をベクトル化したものと 4. で算出

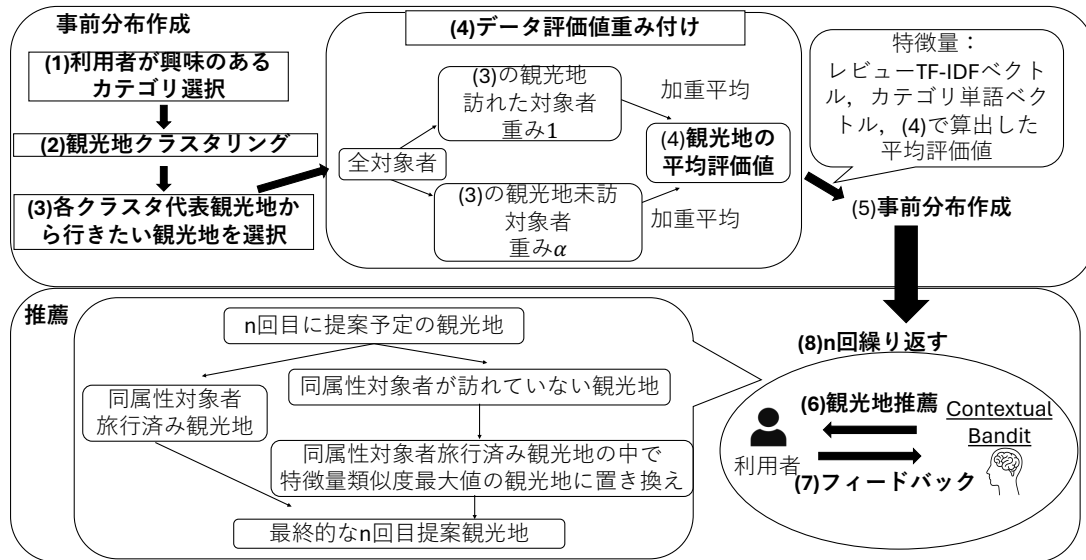


図 4.1 手法手順

した平均評価値，カテゴリの単語ベクトルを特徴量としてどのような観光地が利用者に適合するかを推定するための事前分布を作成する。

## 6. 観光地提案

5. で作成した事前分布を用いた Contextual Bandit アルゴリズムを用いて，観光地を提案する。

## 7. フィードバック

利用者が行きたいと思った観光地をフィードバックし，結果から確率分布を更新する。

## 8. 6. と 7. を交互に $n$ 回行う

$n$  回目推薦において，提示された観光地の中に利用者が選んだ条件に該当する対象者が訪れた記録がない場合，その観光地は利用者の条件に適合していない可能性が高いと考えられる。そのため，特徴量の類似度が最も高い観光地の中から，条件に該当する対象者が実際に訪れた場所を選び代わりに利用者に提案する。

## 4.1 観光地クラスタリングのための特徴量

本節では、クラスタリングで利用する観光地に関する特徴量について説明する。利用する特徴量は、観光地のカテゴリ、観光地に訪れた対象者のレビューをベクトル化したものとする。レビューのベクトル化は TF-IDF を用いる。カテゴリのベクトル化は東北大学の乾・岡崎研究室で作られた日本語 Wikipedia の本文全文を元にした学習済みモデル\*を使用した word2vec[6][5] で行った。観光地のカテゴリは、高級ホテル、温泉、ビーチが例として挙げられる。観光地のカテゴリは、利用者に観光地を提示する際に同じカテゴリの観光地だけ提示しないようにするために用いた。観光地のカテゴリを特徴量として用いることで同じカテゴリの観光地が同一クラスタに分類することが可能であると考えた。また、観光地に訪れた対象者のレビューは、「夜景が綺麗」、「海が絶景」といったものが例として挙げられる。レビューの TF-IDF ベクトルは、コーパス（文書集合）の単語を語彙として定義し、それを次元とする数値ベクトルである。例として、レビューの文書集合があるとし、文書集合中に存在する単語集合  $V = \{ \text{夜景, 綺麗, 温泉, ホテル, 食事} \}$  と「夜景が綺麗だ」というレビューがあるとすると、レビューの TF-IDF ベクトルは  $V$  の TF-IDF 値を  $v = \{0.06, 0.06, 0.00, 0.00, 0.00\}$  とすると 4.1.1 式で表す。

$$\mathbf{v}_{d_1} = [0.06 \quad 0.06 \quad 0.00 \quad 0.00 \quad 0.00] \quad (4.1.1)$$

観光地に訪れた対象者のレビューの TF-IDF ベクトルを用いることで、レビュー中で重要とされる単語が類似している観光地を同一クラスタに属させることが可能であると考えた。

算出した特徴量を用いて  $k$ -means 法を用いて観光地をクラスタリングする。具体的には、算出した特徴量とクラスタの重心の距離を測るために利用する。

## 4.2 選択肢限定のためのクラスタリング

本節では提案手法の手順 2 のクラスタリングする方法について説明する。我々は、提案手法の手順 2 で利用者に観光地を提案する際に 1. で入力してもらったカテ

---

\*[https://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki\\_vector/](https://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/)

ゴリに該当する観光地全てを利用者に提示してしまうと、利用者の選択肢が膨大な数になってしまい、利用者が全ての観光地を調べるための時間がかかってしまうと考えた。さらに、選択肢の数が多すぎると利用者は全ての選択肢を見るために多くの時間を必要としてしまい、利用者は全ての選択肢を見ずに選択してしまう可能性が高く、利用者個人の好みを正確に把握することが困難になってしまうと考えた。そこで、本提案手法では、カテゴリに該当する観光地を 4.1 節で述べた特徴量を用いて  $k$ -means 法でクラスタリングすることにした。本研究では利用者ごとに提示される観光地の数を統一するためクラスタ数を予め決めておくことにした。クラスタリングすることで、我々は利用者に提示する選択肢の数を抑えることができ、利用者が全ての選択肢に目を通した上で選択することが可能であるためより正確に利用者の好みを把握することが可能であると考えた。

### 4.3 対象者の重み付けと平均評価値算出

本節では、手順 4 で述べた評価値の加重平均について説明する。

データに格納されている観光地  $i$  に訪れた対象者  $j$  の  $i$  に対する評価値を  $[0,5]$  で表したものを  $x_{i,j}$ 、 $j$  に付与された重みを  $w_j$  とすると、 $i$  の評価値の平均  $\bar{x}_i$  は (4.3.1) 式によって求められる。ただし、重み  $w_j$  は 4. で付与された対象者は  $w_j = \alpha$ 、4. で重みが付与されていない対象者は  $w_j = 1$  とする。

$$\bar{x}_i = \frac{\sum_{j=1}^n w_j x_{i,j}}{\sum_{i=1}^n w_i} \quad (4.3.1)$$

加重平均の結果を各観光地の新たな評価値とする。対象者の利用者との嗜好の類似度を考慮して重みを付与することで、利用者との嗜好が類似していない対象者の行動履歴データによる影響力を抑えることが可能であると考えた。

### 4.4 作成した事前分布を利用した Contextual Bandit

本節では作成した事前分布を用いて Contextual Bandit を実行する手順を説明する。以下に手順を述べる。



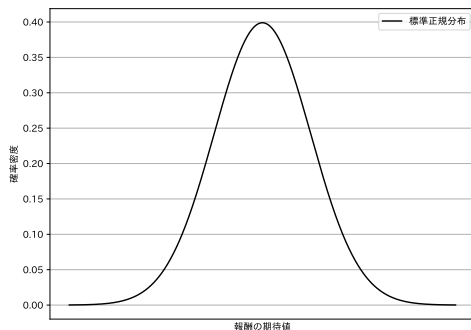


図 4.2 標準正規分布

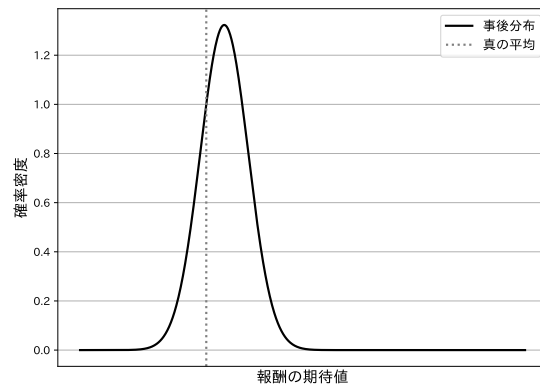


図 4.3 利用者に事前に興味のある観光地を選んでもらって作成した事前分布

1. 4 の手順 4. で算出した平均評価値とレビューの TF-IDF とカテゴリの word2vec を抽出してコンテキスト (特徴量) とする.
2. 4 の手順 3 の結果で利用者に選ばれた観光地には報酬 1 を、選ばれていない観光地には報酬 0 を付与して事前分布を作成する.
3. 手順 2 の事前分布を用いて報酬が最大となるアーム, つまり, 利用者が好むとされる確率が最も高い観光地を Thompson Sampling を用いた Contextual Bandit で割り出す.

#### 4.5 非同属性対象者の行動履歴データを利用者への観光地提案処理

提案手法の手順 7 までは条件に全対象者の行動履歴データを活用していた. しかし, 最終的には利用者が選択した条件に該当する観光地を提案する必要がある. 我々は最終推薦の際に Contextual Bandit で非同属性対象者が訪れた観光地が推薦された場合, 同属性対象者が訪れた観光地に置き換えて利用者に推薦する必要があると考えた. 我々は, 非同属性対象者が訪れた観光地と最も特徴量類似度が高い, 同属性対象者が訪れた観光地を提案した方がより利用者に最適な観光地が提案可能であると考えた. ここで使用する特徴量は, 4 の手順 5 で使用した特徴量と同じである. 提示するものは本提案手法の手順 2 と同様で場所の名前, カテゴリ, 評価,

場所の URL である.

ここで,  $\mathbf{x}$  を同属性対象者の行動履歴データの特徴量ベクトル,  $\mathbf{y}$  を非同属性対象者の行動履歴データの特徴量ベクトルとすると各特徴量のコサイン類似度は以下の計算式で表す.

$$\cos(\mathbf{x}, \mathbf{y}) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \quad (4.5.1)$$

## 第 5 章 評価実験

本実験では提案手法の精度が比較手法よりも高いことを確かめるために、三つの比較手法を用いて比較実験を行った。比較手法 1 は、データを被験者と同属性対象者の行動履歴データのみを用い、非類似嗜好対象者の行動履歴データに重み付けを行わない Contextual Bandit である。比較手法 2 は、被験者と同属性対象者データのみを用いて非類似嗜好対象者の行動履歴データに重み付けをせず、事前分布作成を行わない Contextual Bandit である。比較実験 3 は、ユーザベース協調フィルタリングである。各比較実験の内容と確認することは、以下の通りである。

### 実験 1 提案手法と比較手法 1 の最後の提案による被験者適合率を比較

被験者と非同属性対象者の行動履歴データを活用し、さらに非類似嗜好対象者の行動履歴データに低い重み付けをしたことにより、同属性対象者の行動履歴データのみを使用した場合より被験者の平均適合率が向上していることを確認する。

### 実験 2 比較手法 1 と比較手法 2 の最後の提案による被験者適合率を比較

作成した事前分布を用いた Contextual Bandit が標準正規分布  $N(0, 1)$  を用いた Contextual Bandit よりも利用者の平均適合率が高いことを確認する。

### 実験 3 本提案手法を用いた最後の推薦結果の適合率と比較手法 3 の適合率を比較

提案手法が協調フィルタリングと比べて被験者の平均適合率が高くなっていることを確認する。

## 5.1 データセット

今回使用したデータセットは 4travel\* からスクレイピングしたものである。本実験でスクレイピングした内容は対象者が訪れた観光地の名前、対象者の観光地に対する評価、レビュー、カテゴリ、URL である。本実験では、提案手法と比較手法両方とも同じデータセットを利用する。

---

\*<https://4travel.jp/>

表 5.1 使用データ数

	同属性対象者	非同属性対象者
対象者数	10	12
合計データ数	991	1221

## 5.2 実験手順

提案手法の実験は、4 の手順通りで行う。ただし、クラスタ数  $k = 20$ ，重み  $\alpha = 0.3$ ，Contextual Bandit を用いた観光地提案と被験者からのフィードバックを交互に  $n = 3$  回行うとする。比較手法 1 の実験手順は提案手法の実験手順と同じであるため、省略する。比較手法 2 は、4 の手順 6, 7, 8 を事前分布を  $N(0, 1)$  の標準正規分布にした状態で  $n = 3$  で行う。比較手法 3 の実験手順は以下の通りである。

1. 行きたい場所のカテゴリを被験者が選択する。
2. 1. で選ばれたカテゴリに該当する場所を被験者に提示し、行きたいところを選ぶ。
3. 2. の結果を用いてユーザベース協調フィルタリングを用いて利用者に観光地を提案する。
4. 上位 20 件の結果の適合率を求める。

### 5.2.1 ユーザベース協調フィルタリング

本節では、比較手法 3 で用いるユーザベース協調フィルタリング [13] について説明する。協調フィルタリングの手順は以下の通りである。

1. 被験者と対象者が訪れたら 1，訪れていないならば 0 の評価をつけたものの表を作成する。
2. 被験者と対象者間の類似スコアをコサイン類似度で算出する。
3. 被験者が訪れていない観光地で、類似度が最も高い対象者が訪れた観光地を 20 件推薦する。ただし、被験者が訪れていない観光地で、類似度が最も高い

表 5.2 各実験の適合率と  $t$  検定を行った際の  $p$  値

手法	適合率
本提案手法	0.461
手法 1	0.413
手法 2	0.456
手法 3	0.506

表 5.3 各実験の  $t$  検定を行った際の  $p$  値

手法	$p$ 値
実験 1	0.224
実験 2	0.345
実験 3	0.345

対象者が訪れた観光地が 20 件もない場合は、二番目に類似度の高い対象者が訪れた観光地も提示することとする。

### 5.3 結果・考察

各手法を用いた実験の平均適合率に有意な差があるか確かめるために、 $t$  検定を行った。帰無仮説は提案手法と各比較手法の平均適合率に有意な差はない。である。有意水準は  $\alpha = 0.05$  で、2 標本  $t$  検定を行った。 $p$  値が 0.05 を下回った場合に帰無仮説が棄却され、精度の変化に有意な差があるといえる。各実験の検定方法については実験 1 は対応のない 2 標本  $t$  検定、実験 2 と実験 3 は対応のある 2 標本  $t$  検定である。

表 5.2 に各手法の平均適合率を示し、表 5.3 に各実験の  $t$  検定を行った際の  $p$  値を示す。

適合率を比較すると、どの実験も有意差があることの確認はできなかったが、提案手法の方が適合率が高い被験者が存在した。

実験 1 の結果から、非同属性対象者の行動履歴データも利用し、さらに非類似嗜

好対象者の行動履歴データに低い重みを付与した Contextual Bandit の方が平均適合率が高いことが確認できた。全被験者のうち半数の被験者が、非同属性対象者の行動履歴データも利用した Contextual Bandit の方が適合率が高いことがわかった。しかし、全被験者のうち残り半数の被験者は、同属性対象者データのみを使用した方が適合率が高い、またはどちらも同じ適合率であった。被験者の中には、二回目の試行と三回目の試行の適合率を比べると、三回目の適合率の方が低いという被験者がいた。我々は、三回目の提案で適合率が下がってしまったり、平均適合率に有意な差がなかったりした原因として、提案手法の (8) における最終提案の際に、同属性対象者が訪れていない観光地を同属性対象者が訪れた観光地に置き換えて提案するという方法において、観光地の置き換えが被験者の嗜好に基づいていない可能性があると考えた。

実験 2 の結果から、作成した事前分布を用いた Contextual Bandit の有効性は確認できなかった。しかし、被験者の半数が、作成した事前分布を用いた Contextual Bandit の方が精度が良いという結果が出ているため、被験者によっては事前分布を作成した方が適合率が上がる場合もあることが確認できた。我々は、提案手法の平均適合率が向上しない原因として、作成した事前分布では被験者の嗜好を把握できなかった可能性があると考えた。

我々は、4 の手順 8 における  $n$  回目推薦する観光地置き換えの際に特徴量を利用している。また、我々は事前分布を作成する際にも特徴量を利用している。実験 1 と実験 2 の二つの結果から、提案手法で利用した特徴量が提案手法の平均適合率向上を妨げる原因となっている可能性があると考えられる。我々は、特徴量の一つに観光地の総合評価値を用いていた。観光地の評価値には総合評価値だけでなく、サービスや雰囲気といった各観点における評価値がある。我々は、同じ評価値でも各観点における評価値が異なる場合もあるため、全ての観点における評価値も特徴量に追加すれば、提案手法の適合率が向上するのではないかと考えた。また、もう一つの特徴量としてレビューの TF-IDF ベクトルも利用していた。TF-IDF は単語の文章中における重要度を表すものである。そのため、ある二つのレビューが同じ TF-IDF ベクトル値であっても、必ずしも二つともポジティブまたはネガティブなレビューであるとは限らない。どちらかがポジティブな内容でもう一方がネガティブな内容であるという場合もあり得る。後者の場合、TF-IDF ベクトル値の類

似度が高くても全く評価の異なる観光地であると言える。後者の場合に選ばれた観光地は、被験者にとって最適であるとは言えない。我々は、以上のことを踏まえ、レビューの TF-IDF ベクトルだけでなくレビューの感情スコアも特徴量に追加すれば、さらに類似度の高い観光地を提案できるのではないかと考えた。

実験 3 の結果から、提案手法が利用者と同属性対象者の行動履歴データのみを用いた協調フィルタリングより高い平均適合率を出すことは確認できなかった。しかし、被験者全体のうち半数の被験者が協調フィルタリングより提案手法の方が平均適合率が高いことが確認できた。ユーザベース協調フィルタリングと比べて平均適合率が低くなった原因として、実験 1 と実験 2 の結果をもとに述べたように特徴量に問題点があったと考えられる。また、提案手法の一回目の推薦から三回目の推薦までの適合率の推移を見たとき、どの被験者も適合率多少の上下があっても大きな変化がなかった。そのため、三回の試行だけでは被験者個人の嗜好を十分に学習することができない可能性が考えられる。さらに、実験結果からデータ量が少ない場合におけるユーザベース協調フィルタリングの平均適合率は約半分で、提案手法の平均適合率と 0.45 しか変わらないことがわかった。

全ての実験において有意な差が出なかったため、提案手法の推薦精度は既存手法の推薦精度と変わらないと言える。また、提案手法の平均適合率は 5 割にも達していないため、提案手法の観光地推薦システムは十分な精度であるとは言えない。

## 第6章 おわりに

本研究では、同属性対象者が少ない利用者に対する推薦システムの構築を目指した。本研究では、利用者と同属性対象者の行動履歴データに限定して推薦システムを作るため、全ての対象者の行動履歴データよりデータ数が少なくなってしまう。データ数が少ないと、利用者と嗜好が類似している対象者データはさらに少なくなる。データが少なくなってしまうと、推薦システムの精度が下がってしまうことが本研究の問題点であった。

本研究では、利用者と嗜好が類似していない他の利用者の行動履歴データの影響力を抑えるために0以上1以下の値である $\alpha$ 倍の重み付けすることによってデータ不足による推薦精度の低下を防ぐことを目指す。さらに、我々は利用者が選択した条件に該当しない利用者の中にも利用者と嗜好が似ている人が含まれている可能性が高いと考えたため、利用者の嗜好を把握するために利用者が選択した条件に該当しない利用者データも活用する。また、我々は利用者個人に特化した観光地を推薦するためにContextual Banditと利用者からのフィードバックを用いる。我々は、Contextual Banditを行う際には、事前に利用者にどのような観光地が好みであるのか実際に選んだ結果を利用して最初にどのような特徴量を持つ観光地が利用者に適合している確率が高いのかを示す事前分布を作成することで、より精度の高いContextual Banditアルゴリズムを実行することが可能であると考えた。また、我々は利用者からのフィードバックを受けて確率分布を更新していくことで、より利用者個人に特化した観光地推薦が可能であると考えた。

我々は、提案手法の精度が既存手法より高いことを確かめるために三つの比較手法を用いて実験を行った。一つ目の比較手法は、条件に該当する対象者データのみを使ったContextual Banditである。二つ目の比較手法は、条件に該当する対象者データのみを使って、事前分布を作成せずに標準正規分布を事前分布に用いた場合のContextual Banditである。三つ目の比較手法は、条件を満たしている対象者データのみを使用したユーザベース協調フィルタリングである。

我々は、提案手法と一つ目の比較手法を比べることで、非同属性対象者データを用い、さらに非類似嗜好対象者の行動履歴データに低い重みをつけたことが推薦システムの適合率向上につながるのかどうかを確認した。また、一つ目の比較手法と



二つ目の比較手法を比べることで、事前分布を作成することが推薦システムの精度向上につながるのかどうかを確認した。さらに、提案手法と三つ目の比較手法を比べることで、本提案手法が条件を満たしている対象者データのみを用いたユーザーベース協調フィルタリングより高い精度の推薦ができるかどうかを検証した。

我々は、提案手法と一つ目の比較手法の比較実験で有意差は確認できなかったものの、条件に該当しない対象者のデータも利用した Contextual Banditの方が平均適合率が高いことを確かめることができた。また、一つ目の比較手法と二つ目の比較手法の比較実験で、利用者によっては作成した事前分布を用いた Contextual Banditの方が精度が高くなることを確認することができた。提案手法と三つ目の比較手法の比較実験では、平均適合率では提案手法よりも協調フィルタリングの方が高かったものの、被験者の半数は提案手法の方が適合率が高いことを確かめることができた。

我々は、三つの実験結果から、提案手法で利用した特徴量の一つであるレビューの TF-IDF ベクトルだけでは、対象者の観光地に対する評価がポジティブであるのかネガティブであるのか把握しにくいのではないかと考えた。総合評価値から対象者の観光地に対する評価がポジティブであるのかネガティブであるのかはわかるが、サービスや雰囲気など、どの観点においてどのような評価をしているかはわからない。我々は、特徴量にレビューの感情スコアと観光地の各観点における評価値も加えることで、より利用者に適合する観光地を提案できるのではないかと考えた。

今後の展望として、提案手法の適合率が高くなかったことを踏まえて特徴量にレビューの感情スコアと観光地の各観点における評価値を加えた上で、Contextual Banditの試行回数を増やした場合、協調フィルタリングや他の推薦システムより適合率が上がるのか検証したい。また、データの重み付けを実験では  $\alpha = 0.3$  としたが、 $\alpha$  の値を変化させることによる適合率の変化も確認したい。

## 謝辞

本研究を進めるにあたって、たくさんの方々に支えていただきました。

指導教員である鈴木優准教授には、学会関係や、論文、実験を進めるにあたってたくさんのご指導を賜りました。研究で行き詰まってしまった時にたくさんご相談させていただき本当にありがとうございます。大学院での二年間もどうぞよろしくお願いいたします。

秘書の井尾さんには、様々な手続きを行うにあたって、お世話になりました。おかげさまで、研究活動を滞りなく進めることができましたと思います。ありがとうございます。

鈴木研究室の先輩方には、本当にお世話になりました。研究でわからないことがあったり、論文で行き詰まったり、ポスターの添削をしてもらったりしていただいたおかげでここまでやってこれました。先輩たちがいてくださるおかげで絶対ななんとかなる精神でここまでやってこれました。先輩たちのことを本当に大尊敬しています。本当にありがとうございます。来年からは今の M2 の先輩がいなくなってしまうのでとても寂しいです。あと二年いてほしいけれど無理そうなので諦めます。今の M1 の先輩に絶対に頼りまくりです。大尊敬の川上先輩と高橋先輩よろしくお願ひします。すごく頼りにしています。

同期の皆様にも本当にお世話になりました。研究や本論文を書くにあたって行き詰まっても三人も頑張っているから耐えようと思って頑張れました。田中くんに関しては、院試の志望理由者を書くときに自分の名前を間違えるというミスを犯してくれたおかげで一年間楽しくいじることができました。あと二年精一杯いじります。ありがとうございます。

B3 の皆様、これからよろしくお願ひします。

本論文を無事に書き終えることができたのは、支えてくださった皆様のおかげです。心より感謝申し上げます。

最後に、大学を卒業するまでの間、経済的にも、精神的にも私のことを支えてくれた家族に対しても、感謝の意を表したいと思います。すぐ体調を崩して心配かける娘をこれからもよろしくお願ひします。

## 参考文献

- [1] David Goldberg, David Nichols, Brian M Oki, and Douglas Terry. Using collaborative filtering to weave an information tapestry. In *Proceedings of the ACM Conference on Communications of the ACM*, pp. 61–70. ACM, 1992.
- [2] 平手勇宇. 商品推薦システム—e-commerce サービスでの推薦システムの活用—. 人工知能, Vol. 34, No. 3, pp. 292–299, 2019.
- [3] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. *Advances in neural information processing systems*, Vol. 24, pp. 2249–2257, 2011.
- [4] 本多淳也, 中村篤祥. バンディット問題の理論とアルゴリズム. 講談社, 2016.
- [5] Tomas Mikolov. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, Vol. 3781, , 2013.
- [6] Xin Rong. word2vec parameter learning explained. *arXiv preprint arXiv:1411.2738*, 2014.
- [7] Chris McCormick. Word2vec tutorial—the skip-gram model. *Apr-2016.[Online]. Available: <http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model>*, 2016.
- [8] 天笠俊之, 鈴木優, 宮崎純, 楠和馬. テキストデータマネジメント 前処理から分析へ. 岩波書店, 2022.
- [9] 安江優人, 鈴木優. 情報検索における contextual bandit を用いた適合性フィードバックの利用. In *DEIM Forum 2022 F41-3*, pp. 17–22. 岐阜大学 工学部 電気電子・情報工学科 情報コース 鈴木研究室, 2022.
- [10] 村上知子, 吉岡信和, 折原良平, 古川康一. Cam 法を用いた個人嗜好モデルに基づく商品推薦システム. 人工知能学会論文誌, Vol. 20, No. 5, pp. 346–355, 2005.
- [11] 土方嘉徳. 情報推薦・情報フィルタリングのためのユーザプロファイリング技術. 人工知能, Vol. 19, No. 3, pp. 365–372, 2004.
- [12] 三宅悠介, 峯恒憲. Synapse: 文脈に応じて継続的に推薦手法の選択を最適化する

る推薦システム. 電子情報通信学会論文誌 D, Vol. 103, No. 11, pp. 764–775, 2020.

- [13] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Empirical analysis of predictive algorithms for collaborative filtering. *arXiv preprint arXiv:1301.7363*, 2001.

## 発表リスト

- [1] 尾関采音, 鈴木優『Contextual Bandit を利用した商品レビュー推薦』, 東海関西データベースワークショップ 2024, 2024
- [2] 尾関采音, 鈴木優『Contextual Bandit を用いた特定利用者向け観光推薦システム』, 第 17 回データ工学と情報マネジメントに関するフォーラム, 2025