

行動変容を促すナッジ文言選定のための アプリログによる属性情報の補完手法

吉川裕木子[†] 勝間田優樹[†] 鈴木喬[†] 山田暁[†]
株式会社 NTT ドコモ[†]

1. はじめに

近年、スマートフォンの普及に伴い、個人が持つ端末に情報を送信することで、ユーザ毎に異なる内容を伝達する手法が広く実施されている。また、情報を配信する際に、音・振動・短い文章で知らせることでユーザの興味を惹き、配信した情報の認知率を向上させるプッシュ通知が注目されている。よって、プッシュ通知によるレコメンデーションにおいて、端末を利用しているユーザを分析し、配信内容やタイミングをユーザの興味に沿うように最適化することで、ユーザの情報認知率を最大化させるという営みが様々な企業で行われている。

また、ユーザに情報を提供することが意思決定に与える影響について、行動変容に関する分野の研究が注目されている。ここで述べる行動変容とは、行動経済学でいわれる人間の行動理論に基づいて働きかけることで、その人の行動を変化させることができるというものである。このような働きかけを活用して、環境問題や都市渋滞など様々な社会問題の解決に取り組んだ事例研究がおこなわれている[1][2]。

また、行動変容をさらに促す手法として、ナッジ[3]が注目されている。ナッジとは、人間の思考の癖を利用し、ユーザにとって望ましい選択をするように促す行動変容に関する手法の一つであり、ナッジを活用した研究や実例が多く実施されている。例えば、プッシュ通知によって表示されるメッセージにナッジを適応することでプッシュ通知の開封率を向上させるという効果が挙げられている[4][5]。

ナッジには複数のフレームワークが存在する。例えば、人間の得をすることより損をすることの方を大きく評価するという損失回避バイアスという心理特性を活用した「損失回避フレーム」や、他人と同じようにふるまうことを好むという同調性バイアスという心理特性を活用した「同調フレーム」などが例に挙げられる[6]。

本研究では、ユーザのもつ心理特性によって、異なるフレームワークによって設計されたナッジ種別の効果が異なることに着目し、ユーザの心理特性の度合いを推定することで、ユーザに最も効果のあるナッジの種別を推定することができる。さらに、その推定に基づいて、最適

なナッジを選定し、プッシュ通知に活用することで、プッシュ通知の開封率の向上をはかることができる。しかし、このナッジ種別の推定にはユーザの属性情報が必要であり、その情報量はナッジ種別の推定に精度に大きくかわる。こういったシステムを実現するには、推定精度の担保のため多くの情報を収集する必要があり、その情報収集にかかるコストがシステム化する際の課題となることが予想される。

よって本提案手法では、属性情報を持たないユーザのナッジ種別の推定に必要な情報の拡張手法を提案する。具体的には、属性情報が取得できている一部のユーザを用いて、属性情報を持たないユーザの属性情報を補完する。本提案手法により、すべてのユーザに対して属性情報を獲得することが可能となり、これは最適なナッジ種別の推定に効果的であると言える。

2. 関連研究

ユーザの意思決定の癖である心理バイアスとユーザのパーソナリティの関係性は、過去の研究で示されている。例えば、時間割引率に関するアンケート分析[7]では、年齢や性別、婚姻の有無などが時間割引率に対して影響があることを示した。このような研究から、ユーザが持つ心理バイアス特性を調査できるようにデザインされたアンケートを用いることで、ユーザ自身の属性情報との関係性を分析できることが分かる。

一方、心理バイアスと対応したフレームワークを用いたナッジとユーザのパーソナリティの関係性についても、過去の研究で示されている。例えば、学習を促すメッセージ配信において、事前に測定したユーザの性格特性と、メッセージ配信に用いた 2 種類のナッジに対する反応に相関関係があったという実例が挙げられている[8]。このような研究から、ユーザのもつパーソナリティに応じたナッジを用いることで、行動変容率の向上が期待できることを示している。

これらの従来研究から、ユーザのパーソナリティを用いて学習モデルを構築することで、そのユーザの行動変容を促すのに最適なナッジを推定することが可能だといえる。しかし、こういった学習モデルの構築に必要な、ユーザのパーソナリティは、年収や性別などの属性情報に関するア

Proposal and Evaluation for Nudge Optimization Recommendation System Using Extending Attribute Information with Application Logs
[†]YUKIKOYOSHIKAWA, YUKI KATUMATA, TAKASHI SUZUKI, AKIRA YAMADA, NTT DOCOMO Ltd.

ンケートや、くじの購入金額の質問[9]のようなユーザがもつ心理バイアス特性を特定するようなアンケートを実施することで収集する、といったことが従来の研究では行われてきた。しかし、こういったアンケートの質問設計は大変複雑であり、またユーザから有効な回答が得られる割合多くはない。よって、情報収集の面でナッジのユーザ最適化を実サービスとして実現するのが難しいといえる。

ユーザのパーソナリティを示すデータの収集方法に、ユーザのスマートフォンなどから収集できるログを活用して推定する研究がおこなわれている。例えば、携帯端末から収集できる位置情報などからわかる、日常的な行動パターンから職業属性が推定できる [10]。また、Web サイトにおける時系列的なアクセス数から、年齢などのユーザ属性に基づいたクラスターを推定することができる[11]。

これらのことから、パーソナリティデータを収集するために、スマートフォンから自動的に収集できるログを活用することで、アンケートなどから収集できる情報相当を推測できると考えられる。

よって、本研究では、一部のユーザに実施したアンケートなどから収集した情報と、全てのユーザから自動的に取得可能なスマートフォンからのログを用いて、ナッジのユーザ最適化に有効な属性情報を補完する手法を提案する。また、ナッジのユーザ最適化において本提案手法が有効であることを示す。

3. 提案手法

3.1 ユーザが持つ情報種別について

本研究では、図 1 のように 2 種類のユーザ群を対象ユーザとする。ユーザ群 A は属性情報とアプリログの 2 種類のデータを持ち、ユーザ群 B はアプリログのみをもつ。また、属性情報とアプリログはベクトル表現に変換されており、ベクトル間距離を算出することで、ユーザ間の類似度を示すことができる。

ここで属性情報は、ユーザ群 A の過去のアンケート回答や他のサービス利用から推測される、性年代や趣味趣向のことを指し、アプリログは本システムが導入されたアプリケーションから取得できる、各画面の閲覧時間やボタン押下回数などの操作ログを指す。

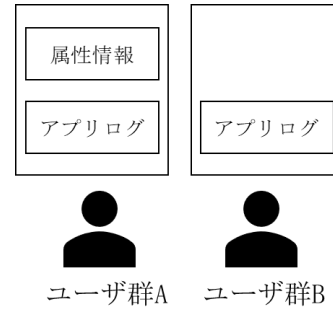


図 1 本研究で用いたユーザ群と保持するデータ

3.2 アプリログのベクトル空間距離の算出

2 で述べたスマートフォンからパーソナリティを推定できるという従来研究より、本研究における属性情報とアプリログには相関関係があり、アプリログが類似しているユーザ同士は属性情報も類似していると考えられる。本提案手法では属性情報とアプリログの類似度をベクトル間距離と定義する。

まず、属性情報とアプリログのベクトル間距離を定義する。ここで、ユーザ i とユーザ j 間の属性情報 X_d の距離は以下の式で与えられる。

$$distance_{Xd} = \sqrt{(x_i^d - x_j^d)^2} \tag{1}$$

また、アプリログ X_a の距離を以下の式で与える。

$$distance = \sqrt{(x_i^a - x_j^a)^T W (x_i^a - x_j^a)} \tag{2}$$

$$W = \begin{bmatrix} W_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & W_n \end{bmatrix} \tag{3}$$

ここで、 W はアプリログの重み付けパラメータであり、アプリログの距離が近いと属性情報の距離が近いという関係が成り立っている。なお、パラメータ W の学習は以下の式を用いて、対象ユーザの全通りの組み合わせで行われる。

$$Loss Function = distance_{Xd} - distance \tag{4}$$

$$W' = W + \frac{\partial Loss}{\partial W} \tag{5}$$

次に、属性情報とアプリログの両方をもつユーザ群 A を用いて、パラメータ W を学習する。式(2)に学習済みの W を用いて、アプリログのユーザ間距離が近いユーザは同じクラスターになるようクラスタリングを行い、各クラスターに所属するユーザの平均ベクトルを獲得する。

3.3 属性情報の補完手法

アプリログのみをもつユーザ群 B のユーザ B の属性情報を補完する手法を示す。ユーザ B のアプリログと 3.2 で獲得したクラスターの平均ベクトルとの距離を、式(2)を用いて

算出し、ユーザ B に最も近いクラスタを選択する。

次に、選択されたクラスタ内からユーザ群 A に所属するユーザ A をランダムで抽出し、そのユーザがもつ属性情報をユーザ B の属性情報とする。なお、ランダムで選ぶ理由として、ユーザ群 A とユーザ群 B にアプリログの傾向が異なる可能性が考えられるため、最も近いユーザを選んだ場合、常に同じユーザが選ばれてしまうことを避けるためである。



図 2 「まくはり MaaS」の画面イメージ

4. 評価手法

4.1 実証実験

本提案手法の評価は、2022年2月1日から2022年3月21日の期間で実施された、幕張新都心エリアにおける実証実験で収集されたデータを用いて行った*1。本実証実験は、地域住民の回遊性向上を目的とした実証実験であり、実証参加者にはアプリケーション「まくはり MaaS」をインストールしてもらい、それを活用しながら幕張新都心エリアを周遊する。図 2 のように、アプリケーションからは、エリアを周遊するための交通手段の予約や、周遊エリアのスポット情報にアクセスすることができる。また、本評価手法で用いるユーザ数の内訳を表 1 に示す。

表 1 評価で用いるユーザ数内訳

	総ユーザ数	1 回以上開封有のユーザ数			
		文言 1	文言 2	文言 3	合計
ユーザ群 A	835	118	119	112	245
ユーザ群 B	3125	345	378	366	773

4.2 プッシュ通知について

実証実験で用いたプッシュ通知システムについて説明する。本システムでは、ユーザの状況をトリガとして、店舗と文言を該当のユーザへプッシュ通知を行う。ユーザの状況とは、アプリケーションをインストールした端末から取得できる位置情報や、特定の時刻、交通手段の予約履歴などのことを示す。ユーザ状況をトリガにしたプッシュ通知は、「幕張駅に着いた時に、周遊スポットをおすすめ」や、「お昼時に飲食店をおすすめ」といった内容のレコメンドをユーザに提供することが可能である。

また、メッセージ確認画面は図 3 のようになっている。プッシュ受信時には、通知タイトルに「あなたへのおすすめがあります」、通知文に「あなたにおすすめの店舗でお食事はいかがですか?」のプッシュ通知が通知エリアに表示される。なお、通知文にはプレーン・損失回避フレーム・同調フレームを活用して設計された 3 種類のナッジ文言を採用した。



図 3 メッセージ確認画面

表 2 アプリログの例

プッシュ開封時間	カーシェア利用回数
アプリ起動回数	シェアサイクリング利用回数
プッシュ開封率	タクシー配車サービス利用回数
クーポン利用回数	バス利用回数
メッセージ詳細閲覧回数	バス予約時間/乗車時間
スポット詳細閲覧回数	ルート検索条件(歩く速度・並び替え)
スポット検索回数	ルート検索使用回数

4.3 最適なナッジ文言の推定

本研究の評価で利用する実証実験の開封履歴について、実証実験中に介入があったことを説明する。具体的には、本実証実験において、ユーザ群 A に所属するユーザの半数には属性情報に応じて推定された文言のプッシュ通知を、もう一方にはランダムで配信を行った。なお、ユーザ群 B に対しては、すべてのユーザに対してランダムで配信した。

ユーザ毎の最適なナッジ文言の推定には、文脈付きバンディットアルゴリズム[12]を採用し、属性情報と過去の開封履歴から学習を行った。バンディットアルゴリズムは、

*1 https://www.docomo.ne.jp/binary/pdf/info/news_release/topics_220125_00.pdf

複数の選択肢から利得を最大にするのを目的としており、新しい情報の探索と、既存情報に基づく選択肢の実行を同時に実施するものである。本実証実験では、選択肢を3種類の文言、情報の探索と実行をプッシュ通知、利得をその開封と定義して学習を行う。具体的には、過去の開封実績とユーザの属性情報に基づいて、文言のユーザ毎評価値を算出する3つのパラメータの更新を夜間に行い、日中はトリガが発火した際に、その評価値に基づいた確率でプッシュ通知を実施するものである。また、ユーザ群Aにおける属性情報に応じたナッジ文言の推定による介入効果を検証したところ、有意差10%で優位に開封率が向上したことがわかった。

4.4 属性情報とアプリログ

実証実験の参加者から、事前同意を得て取得した属性情報とアプリログについて説明をする。属性情報は、NTTドコモが過去に実施したアンケートなどから収集したデータであり、性年代や職業といったものが挙げられる。このような情報を、141次元のベクトル表現にしたものを用いる。一方、アプリログは、4.1で紹介したアプリケーションから特徴量を作成した。アプリケーションは、各種モバイルサービスの利用ボタンや、経路検索機能、クーポン利用画面を所持しており、ユーザによって操作の差異が現れると考える。表2のような情報を、44次元のベクトル表現にしたものを用いる。

4.5 プッシュ通知開封の推定による評価

本提案手法の評価は、実証実験の開封履歴を用いたオフライン検証によって行う。また、本研究の目的であるユーザ毎の最適なナッジ文言の推定を、ユーザ毎の各文言のプッシュ通知の開封を予測する問題として定義し、本提案手法の評価を行う。この予測問題に対して、PythonのオープンソースライブラリであるPyCaret[13]によって、ナッジ文言毎の機械学習モデルを構築する。また、機械学習モデルの説明変数として、アプリログと3.3で補完した属性情報を用い、目的変数としてプッシュ通知に対する開封を用いる。

本研究における評価値の指標として、PyCaretの機能によって選定された最も性能のよいモデルによる予測値と、正解値との相関値を評価値とする。

4.6 プッシュ通知の開封の算出について

目的変数に、ユーザ毎に本実証実験期間での合計開封回数を用いることで、予測値を算出することができる。なお、プッシュ通知回数が極端に少ないユーザや、開封自体が少ないといったユーザが存在するため、開封毎に何回目のプッシュ通知かを考慮した重み付けを行う。ここで、図4に本実証実験のプッシュ通知の結果を集計した結果を示す。

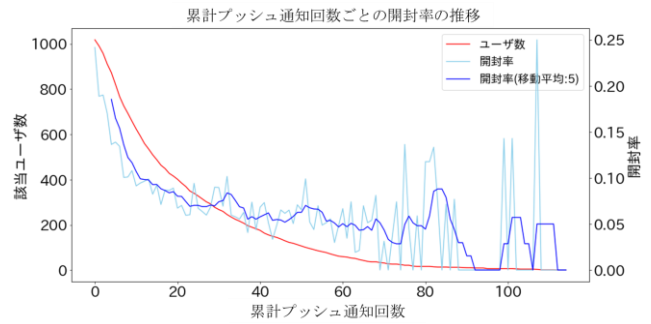


図4 累計プッシュ通知回数ごとの開封率の推移

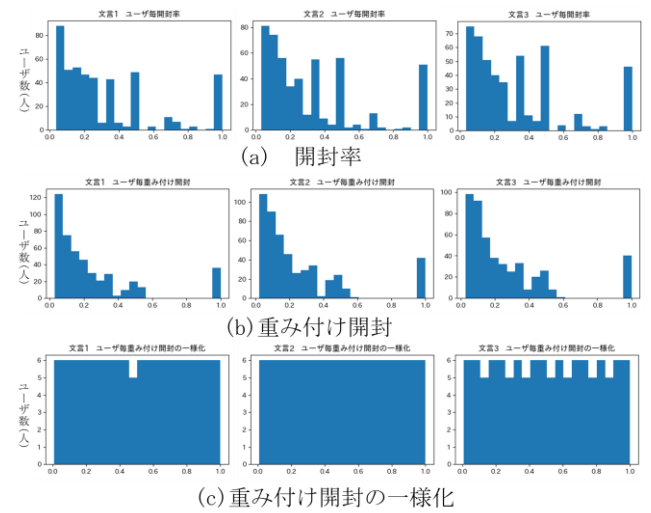


図5 重み付けによる開封Yの変化

図4で示すように、プッシュ通知回数が20回以下の場合、プッシュ通知の開封率が減少傾向にあるが、20回を超えると定常化している。これは、プッシュ通知が少ない間は配信に興味を惹かれ開封するが、何度も配信されるとプッシュ通知に対する興味が低下し、反応しなくなるという事が言える。一方、配信回数に関わらず開封し続けるユーザが一定数いることも示している。

累計プッシュ通知回数が少ないときより、多いときに開封したものが、1回の開封の価値が大きくなるような重み付けを図5(a)に行った結果を図5(b)に示す。また、重み付けは以下の式で行う。

$$Y^{j,a} = \frac{1}{n^{j,a}} \sum_{i:\pi_i^a=j} \frac{\log(1/p_i)}{\sum_{k:\pi_k^a=j} \log(1/p_k)} \quad (6)$$

ここで、Yはユーザaの文言jの重み付け開封である。nはユーザaの文言jの開封回数を、Pは全ユーザのk回目の平均開封率の移動平均を、 $\pi_i^a=j$ はユーザaのi回目のプッシュ受信の文言が文言jであることを示す。なお、60回以上のユーザ数が少ないことから、60回以上のプッシュ通知に関しては無効とする。さらに、重み付け開封Yが一様分布になるように、Yの大きさの順序を合計ユーザ数で割ったY'を予測値として用いる。一様化を行った結果を図5(c)

に示す。

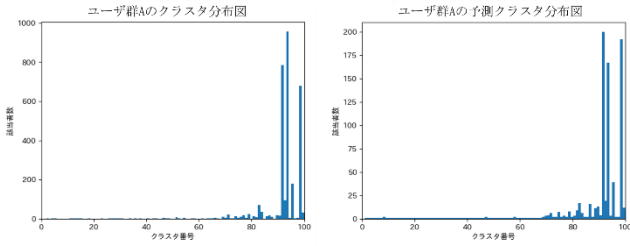


図 6 ユーザ群 A の予測クラスタ分布

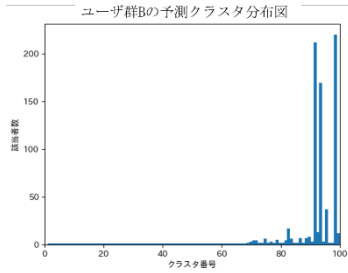


図 7 ユーザ群 B の予測クラスタ分布

	ユーザ群 A	ユーザ群 B
ユーザ数	835	3167
ユーザ所属クラスタ数	100	65
クラスタ所属平均ユーザ数	8.35	48.07
標準偏差	34.431	172.114
変動係数	4.123	3.580

表 3 ユーザ群 A とユーザ群 B の予測クラスタ結果

5. 評価結果

5.1 ユーザのクラスタリングに関する評価

ユーザ群 A に所属する 835 人のユーザの 141 次元の属性情報ベクトルと 44 次元のアプリログベクトルを用いて、式(2)のパラメータ W の学習を行った。学習した W を用いて 100 個のクラスタにユーザ群 A をクラスタリングした結果とユーザ群 A の所属クラスタを予測した結果を図 6 に示す。本提案手法による所属クラスタの予測正解率はマイクロ正解率で 94.7%であった。しかし、図 6 のクラスタ分布図からわかるように、多くのユーザは 4 つのクラスタに所属しているため、正解率が高くなっているともいえる。

次に、ユーザ群 B に所属するユーザ 3125 人を用いて、所属クラスタの予測を行った結果を図 7 に示す。また、表 3 に示すように、ユーザ群 A とユーザ群 B のクラスタリング結果の比較をしたところ、ユーザ群 A とユーザ群 B の傾向に大きく乖離がないところから、3.3 の提案手法による所属クラスタの予測が、ある程度可能であるといえる。

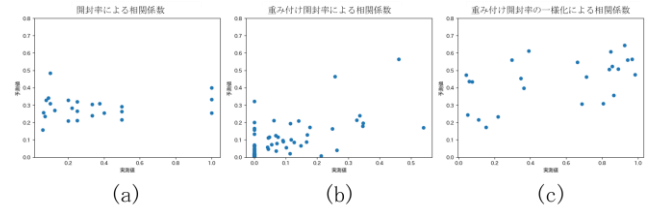


図 8 重み付けによる予測値 Y の分散変化

表 4 ユーザ群 A : 学習と評価に用いたユーザ数

	文言 1	文言 2	文言 3
ユーザ数(人)	118	119	112
学習データ	94	95	89
評価データ	24	24	23

表 5 ユーザ群 A : 属性情報による予測結果

	文言 1	文言 2	文言 3
Best model	K Neighbors Regressor	Random Forest Regressor	Ridge Regression
Correlation	0.289	0.123	0.211
MSE	0.120	0.090	0.129

表 6 ユーザ群 A : 属性情報とアプリログによる予測結果

	文言 1	文言 2	文言 3
Best model	K Neighbors Regressor	AdaBoost Regressor	Ridge Regression
Correlation	0.502	0.290	0.339
MSE	0.101	0.082	0.121

5.2 ユーザ群 A に関する評価

4.6 で示した開封に対する重み付けの処理による効果を、ユーザ群 A を用いた文言 1 における開封 Y の予測結果によって示す。予測に用いたユーザは、1 回以上開封を行ったユーザとする。これは、表 1 で示すように、プッシュ通知を一度も開封していないユーザの割合が多いため、全ユーザを用いて学習をしてしまうと、開封 Y の予測が難しくなるためである。説明変数として、属性情報とアプリログを用い、アプリログは開封率と強い相関のある特徴量を除いた 22 次元で推測を行った。

重み付けと一様化を目的変数である開封率に実施し、学習した場合の予測結果を図 8 に示す。予測値が低い値に集中することなく予測が行えていることが分かる。

次に、属性情報とアプリログが、ナッジ種別推定に有効であることを示すために、ユーザ群 A の属性情報のみを用いた場合と、属性情報とアプリログの両方を用いた場合の予測を行った。表 4 に示すように、モデルの学習と評価に用いたユーザは、各文言プッシュ通知を 1 度でも開封したことのあるユーザのみで行い、学習データと評価データは 8:2 で分割した。

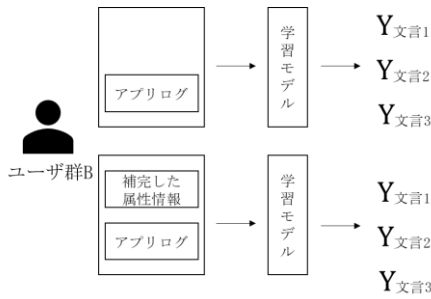


図9 ユーザ B 群を用いた提案手法の評価

表7 ユーザ群 B : 学習と評価に用いたユーザ数

	文言 1	文言 2	文言 3
ユーザ数(人)	345	378	366
学習データ	276	302	292
評価データ	69	76	74

表8 ユーザ群 B : アプリログによる予測結果

	文言 1	文言 2	文言 3
Best model	Random Forest Regressor	Ridge Regression	Ridge Regression
Correlation	0.310	0.158	0.270
MSE	0.07	0.08	0.07

表9 ユーザ群 B : 属性情報とアプリログによる予測結果

	文言 1	文言 2	文言 3
Best model	AdaBoost Regressor	Random Forest Regressor	Random Forest Regressor
Correlation	0.312	0.196	0.379
MSE	0.08	0.08	0.07

属性情報のみを用いた場合と、属性情報とアプリログを用いた場合の予測結果を表5と表6に示す。全ての文言において、属性情報とアプリログを用いた場合が、属性情報のみを用いた場合の相関値を上回っているため、ナッジの開封の予測には属性情報だけでなくアプリログも有効であることが分かる。

5.3 ユーザ群 B に関する評価

本提案手法の有効性の評価方法の概要を図9で示す。ユーザ群 B のアプリログのみを用いた場合と、3で示した提案手法によって選定した属性情報とアプリログの両方を用いた場合の予測結果を比較することで有効性を示す。なお、予測に用いたユーザは、1回以上開封を行ったユーザであり、モデルの学習と評価に用いたデータは表7に示す。表8と表9が示すように、2つの場合を比較した結果、提案手法によって選定した属性情報を用いた場合、予測精度が改善したことがわかる。なお、文言によっては相関値の変化が小さい文言もあり、すべての文言に対して予測精度を

改善するためには、アプリログのベクトルの次元数をより増やすことで表現力を向上させる必要がある。

6. おわりに

本研究では、アプリログを用いて属性情報を補完する手法を提案と評価を行った。ナッジ毎の開封予測精度が向上したことから、本提案手法が有効であるという示唆を得られた。しかし、本研究で行った評価は、過去の実証実験のデータを用いたナッジ毎のプッシュ通知開封の予測によるものであり、本研究の目的であるナッジ種別の推定への有効性を100%示せるものではない。よって、各モデルが出した評価値の違いに基づいて、実際に文言を出し分けることで、その有効性を検証する必要がある。なお、評価値の違いに基づいてナッジの出し分けを行う点については、4.3でユーザ群 A において効果があったことを述べている。今後は、本研究で未検証である内容について実証実験を実施することで、提案手法によるプッシュ通知の介入効果の向上の有効性を検証していきたい。

参考文献

- [1] Reddy, S. M.W., Montambault, J., Masuda, Y. J., Keenan, E., Bulter, W., Fisher, J. R. B., Asah, S. T., and Gneezy, A. "Advancing Conservation by Understanding and Influencing Human Behavior", *Conversation Letters*, Vol.10, No.2, pp.248-256 (2017)
- [2] Xu, W., Kuriki, Y., Sato, T., Taya, M., and Ono, C. "Does Traffic Information Provided by Smartphones Increase Detour Behavior? An Examination of Emotional Persuasive Strategy by Longitudinal Online Surveys and Location Information", *Proc. 15th international conference on Persuasive Technology*, pp.45-57 (2020)
- [3] Leonard, T. C.: Richard H. Thaler, Cass R. Sunstein, "Nudge: Improving decisions about health, wealth, and happiness" (2008)
- [4] Hyun Bae Kim, Toshiya Iwamatsu, Ken-ichiro Nishio, Hidenori Komatsu, Toshihiro Mukai, Yoko Odate, Masanobu Sasaki, "Field experiment of smartphone-based energy efficiency services for households: Impact of advice through push notifications", *Energy and Buildings*, Vol. 223 (2020)
- [5] Carmina G. Valle, Brooke T. Nezami, Deborah F. Tate, "Designing in-app messages to nudge behavior change: Lessons learned from a weight management app for young adults", *Organizational Behavior and Human Decision Processes*, Vol. 161, pp. 95-101, (2020)
- [6] 竹内幹, "社会保障のための行動経済学: 補正か誘導か?" *社会保障研究*, Vol. 6, No. 3, pp. 233-244 (2021)
- [7] 池田新介・大竹文雄・筒井義郎, "時間割引率: 経済実験とアンケートによる分析", *淡路島行動経済学ワークショップ提出論文* (2005)
- [8] Yamauchi Taisei, Takami Kyosuke, Flanagan. Brendan, Ogata Hiroaki, "Nudge Messages for E-Learning Engagement and Student's Personality Traits: Effects and Implication for Personalization", *30th International Conference on Computers in Education Conference, Proceedings* (2022)
- [9] 池田新介, 筒井義郎, "アンケート調査と経済実験による危険回避度と時間割引率の解明", *大阪大学社会経済研究所*,
- [10] 佐藤一夫, "プライバシー保護を考慮したケータイ行動ログの活用について", *情報処理*, vol. 50, No. 7 (2009)
- [11] 佐藤哲, "アクセスパターンのクラスタリングによる Web ログからのユーザ属性推定", *情報科学技術フォーラム* (2016)
- [12] Shipra Agrawal, Navin Goyal, "Thompson Sampling for Contextual Bandits with Linear Payoffs", *the 30th International Conference on Machine Learning*, vol. 28, pp. 127-135 (2013)
- [13] <https://github.com/pycaret>