情報処理学会論文誌 Vol.66 No.2 426–439 (Feb. 2025)

施策効果の高い顧客グループの発見を目的とした機械学習に基づく実験計画手法

中村 友香[1,a)] 山極 綾子[1,b)] 佐々木 北都[2,c)] 後藤 正幸[1,d)]

受付日 2024年4月2日, 採録日 2024年11月1日

概要：EC サイト上でビジネスを行う多くの企業にとって，クーポン配布などのマーケティング施策は一般的になっており，より効率的に行う方法に関する研究がさかんに行われている．これらの施策はユーザの離反を防ぐとともに定着化を図ることを目的に行われるが，全ユーザに一様に施策を実施することは必ずしも得策ではない．施策による影響の大小は個人差があるため，効率的かつ効果的に施策を講じるためには施策によって Life Time Value（LTV）の向上が見込まれるような施策効果が高い顧客群を特定することが重要である．本研究では，ファッション通販サイト ZOZOTOWN 内で古着を販売する ZOZOUSED を対象事例とし，ポイント付与などの施策を講じる対象ユーザの選定に関する問題を考える．これまで ZOZOUSED では，独自のロジックに基づき担当者が自らの知見から施策対象のユーザグループを選定していたが，大量の未開拓のユーザ群の中に高い効果が望めるユーザ群が存在する可能性がある．しかし，大多数の未開拓ユーザの中から効果の高いユーザを発見するための施策実験を行おうとした場合，実験的に行う施策対象者数にはコスト面での上限があり，加えて施策対象者に施策効果が低いユーザを多く含んでしまうことによる利益減少リスクもともなう．そこで本研究では，未開拓のユーザ群の中から高い効果が望めるユーザ群を特定するための定量的指標に基づいた実験計画手法を提案する．具体的には，予測購買単価と予測利用回数の切り口でユーザをセグメント化し，各セグメントにおける施策効果の平均と分散を基に施策対象とするセグメントを決定することで，より効率的な高効果ユーザ群の特定を目指した分析アプローチを構築する．最後に，ZOZOUSED で行った検証実験を基に提案手法の有用性を示す．

キーワード：実験計画，機械学習，効果検証，施策効果，E コマースマーケティング

Experimental Design Based on Machine Learning for Finding Customer Groups with High Measure Effects

Yuka Nakamura[1,a)] Ayako Yamagiwa[1,b)] Hokuto Sasaki[2,c)] Masayuki Goto[1,d)]

Received: April 2, 2024, Accepted: November 1, 2024

Abstract: For many companies doing business on e-commerce sites, marketing measures such as coupon distribution have become common, and research on more efficient methods has been active. However, it is not necessarily advisable to implement measures such as coupon distribution uniformly to all users. Since the impact of a measure differs depending on the individual, it is important to identify a group of customers for whom the measure is highly effective in order to implement the measure efficiently and effectively. In this study, we consider the problem of target user selection for the point award policy as a promotion measure, focusing on ZOZOUSED, which sells used clothes on the fashion shopping site ZOZOTOWN, as the target case study. Currently, as with most e-commerce sites, most of the target users are in the reserve group for defection. There is a possibility that there are unexplored user groups that can be highly effective. Therefore, this study considers the design of a measure experiment to discover a group of users with high measure effectiveness among such a large number of unexplored users. However, in conducting a verification experiment, there is a risk of including many target users with low effectiveness and of restricting the number of target users. In particular, there is a possibility that an unexploited user group exist among the non-respondents of the measures that could be effective in improving its' effect. Therefore, this study proposes an experimental design method based on quantitative indicators estimated by a machine learning model for identifying expected high-effective user groups from among undeveloped user groups.

Keywords: experimental design, machine learning, effect verification, effect of measures, e-commerce marketing

1. はじめに

ECサイト上でビジネスを行う多くの企業にとって，クーポン配布などのマーケティング施策は一般的になっており，より施策が合致する顧客を抽出するなど，それら施策を効率的に行う方法に関する研究もさかんに行われている [1]．本研究では，ファッション通販サイト ZOZOTOWN 内で古着を販売する ZOZOUSED [2] を対象事例とし，ポイント付与などの施策を講じる対象ユーザの選定に関する問題を考える．ZOZOUSED では，ユーザの離反を防ぐとともに定着化を図るためポイント付与などの施策を講じている．この施策では施策対象者数に応じたコストがかかるうえ，施策による影響の大小は個人差があるため，効率的かつ効果的に施策を講じるためには施策によって Life Time Value（LTV）[3] の向上が見込まれるような施策効果が高い顧客群を特定することが重要である．これまで ZOZOUSED では，独自のロジックに基づき施策対象のユーザグループを選定していた．このような経験則に基づいた選定では施策対象が属人的な判断に依存してしまい，施策を適用することでその後の購買頻度が向上するようなユーザに対して適切にコミュニケーションがとれない恐れや，施策効果がほぼないユーザに対する無駄打ちを行ってしまう可能性がある．実際に ZOZOUSED が過去に実施した施策においては，ポイントを付与したユーザが今後の利用が見込めなさそうな離反予備群に偏っているケースもあった．このような状況に鑑みると，施策非対象者中に定着率向上の効果が望めるような未開拓のユーザ群が存在する可能性が考えられる．しかし，大量の未開拓のユーザ群の中から高効果が望めるユーザ群を発見することを目的として施策を実施する場合，施策対象者数の制約に加え，施策対象者に施策効果が低いユーザを多く含んでしまうことによる利益減少リスクをともなう．

本研究では，施策対象者が大きく偏った過去の施策実施データを対象として，施策効果が未知な顧客群の中から施策効果が高い顧客群を特定するための実験計画手法の提案を目的とする．具体的には，未開拓のユーザ群の中から高い効果が望めるユーザ群を特定するための定量的指標に基づいた実験計画手法を提案する．まずは，ユーザを適切にセグメント化したうえで，因果推論を用いた正確な効果推定が難しいユーザ群の中から効果がありそうな群に対して施策実験を設計することを考える．そのために，予測購買単価と予測利用回数の切り口でユーザをセグメント化し，各セグメントにおける施策効果の平均と分散を基に施策対象とするセグメントを決定することで，より効率的な高効果ユーザ群の特定を目指す．

提案手法により，どのような施策実施データに対しても効率的に高効果ユーザ群を特定することが可能となる．これによって，施策効果の高いユーザ群の展開が期待される．

[1] 早稲田大学
Waseda University, Shinjuku, Tokyo 169–8555, Japan
[2] 株式会社 ZOZO
ZOZO, Inc., Chiba 261–7116, Japan
a) krbs_1121@suou.waseda.jp
b) saxophone_0105@ruri.waseda.jp
c) hokuto.sasaki@zozo.com
d) masagoto@waseda.jp

2. 準備

2.1 対象問題

本研究では，ファッション通販サイト ZOZOTOWN 内で古着を販売する ZOZOUSED を対象事例とし，ポイント付与施策を講じる対象ユーザの選定に関する問題を考える．ZOZOUSED では，ユーザの離反を防ぐとともに定着化を図ることを目的として様々な施策を講じている．現状，施策対象ユーザは独自のロジックに基づいて選定されているが，ZOZOUSED の過去実施した施策においては，ポイントを付与したユーザが離反予備群に偏っているケースもあった．ここで問題点が 2 つあげられる．1 つは，現状のロジックにおいて対象にはあてはまらないユーザ群の中に施策効果が高い顧客が存在する可能性が考慮されていない点である．もう 1 つは，顧客が極端に偏ったデータを用いて効果推定を行う場合，施策効果を正しく推定できない顧客群が多く存在する点である．上記の問題点を解決するためには，すべての顧客に対し施策実験を行う必要があるが，この場合コストの大幅な増加が懸念される．したがって，本研究では，実験対象として施策効果が未知の顧客群の中から効果が高そうな顧客群を選定する手法について検討する．

2.2 処置効果の定義

施策効果は一般に，似たようなユーザ群内における施策を実施する群（処置群）と施策を実施しない群（対照群）の群間の結果に関する平均値の差分で定義される．本問題設定では，施策の実施による LTV の向上が見込める顧客群を選定することが求められるため，結果は施策前後における LTV の増加分と定義することが妥当である．ここで LTV とは，1 人，あるいは 1 社の顧客が，特定の企業やブランドと取引を始めてから終わりまでの期間（顧客ライフサイクル）内にどれだけの利益をもたらすのかを算出したものである．この定義に準ずると ZOZOUSED における LTV は顧客の総購入額となるが，商品の金額のばらつきが大きいがゆえに，そもそも効果を算出する以前に購入金額を推定することが困難である．また，ZOZOUSED での施策実施の目的は顧客の定着化にあるため，購入金額の増加より利用回数の増加の方が効果として望ましい．したがって，本研究では，処置群と対照群の群間の施策前後における利用回数の増加分の平均値の差分を施策効果と定義する．**図 1** に効果のイメージ図を示す．この例では，施策効果は 3 回

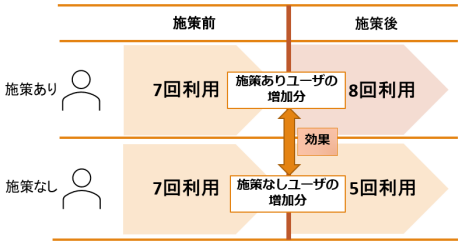


図 1 処置効果のイメージ図
Fig. 1 The image of treatment effect.

となる．

2.3 統計的因果推論と施策効果の推定

施策効果の大きさは，同様のユーザに対して施策を講じなかった場合の LTV と，講じた場合の LTV の差分によって測ることができる．しかし，同一ユーザに対して両者の場合の LTV を同時に観測することは不可能である．そこで，そのような場合でも疑似的に効果を算出する手法として統計的因果推論の枠組み [4] がある．

因果推論は何らかの処置 $t = \{0, 1\}$ が結果 Y に及ぼす効果を推測するアプローチである．通常，私たちは同一個体に対して処置を施した場合の結果 Y_1 と，施してない場合の結果 Y_0 を同時に観測することはできない．そのため，以下の「強く無視できる割当て条件」を仮定し，処置効果を推定するというアプローチをとる．

＜強く無視できる割当て条件＞

$$(Y_0, Y_1) \perp t \mid \boldsymbol{x}, \text{ and } 0 < p(t = 1 \mid \boldsymbol{x}) < 1 \quad \text{for all } \boldsymbol{x} \tag{1}$$

処置以外で結果に影響を与える特徴量である共変量 $\boldsymbol{x}$ を所与とするとき「強く無視できる割当て」であるとは，「処置が施されるかどうかかを表す割当ては観測された共変量の値に依存し，観測結果の値の高低によっては依存しない」という条件のことである [1]．割当ては結果の測定より時間的に先行しているため，結果の値によって割当てが決まるということはほとんどありえない．したがって，この条件の重要な点は「割当て時までに観測された共変量」によって割当てを説明できなくてはならず，観測されていない共変量が割当てに影響を与えていないという点である．この条件を満たすとき，処置が施される確率は共変量によって推定でき，この確率を傾向スコアと呼ぶ．傾向スコアが近い個体どうしの比較は共変量の割当てへの影響を無視できるため，1 個体の処置効果を算出する際などに用いられる．また，処置効果は次式で表される．

(1) 平均処置効果（ATE）

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}[\tau(\boldsymbol{x})] &= \mathbb{E}[Y_1 - Y_0] \\ &= \mathbb{E}[Y_1 \mid t = 1] - \mathbb{E}[Y_0 \mid t = 0]\end{aligned} \tag{2}$$

(2) 個別処置効果（ITE）

$$\begin{aligned}\tau(\boldsymbol{x}) &= \mathbb{E}[Y_1 - Y_0 \mid \boldsymbol{x}] \\ &= \mathbb{E}[Y_1 \mid \boldsymbol{x}, t = 1] - \mathbb{E}[Y_0 \mid \boldsymbol{x}, t = 0]\end{aligned} \tag{3}$$

このように，ある対象者の処置効果 $Y_1 - Y_0$ の直接推定が困難な場合においても，処置 t ごとの結果の期待値を用いることで期待処置効果が推定できる．しかし因果推論は観察データを基に効果を推定する手法であるため，未知のユーザ群に対する効果推定を行うことは難しい．

施策実験を行ううえで，最終的に得られた結果から各ユーザの施策効果を高い精度で推定できることが望まれる．そのための実験計画手法として，ランダム化比較試験 [5] があげられる．これは，ユーザをランダムに施策対象と非対象に分ける手法である．ランダム化することで，検証したい効果以外の要因がバランスよく分かれるため，公平に比較することが可能となる．しかし本問題設定においてこの手法を用いた場合，実験が大規模になりすぎてしまううえ，今までに得られた観察データにおける施策効果に関する情報の活用が困難となる．

したがって，本問題設定では，観察データを活用して高い施策効果が得られそうな施策実験対象者を適切に定義することが求められる．

3. 関連研究

本研究では実際の企業において施策効果の高い顧客グループを特定するための施策実験計画の方法論の提案を行う．本章では，本研究に関連が深い従来研究として，施策実施による影響に対する研究ならびに本研究対象であるアパレルサイトを対象とした研究について概要を示す．

3.1 施策実施に関する研究

施策実施効果を定量化しようとする研究は様々行われている [6], [7], [8], [9]．Narita ら [6] は，期待報酬の推定量に着眼し，期待報酬を推定するためのバッチデータの最も統計的に効率的な利用法についての研究を行っている．また，Kallus ら [7] のオフライン評価に関する研究では，複数の施策のログデータを用いてオフライン評価を行う際に生じる推定量の分散の施策データに対する依存性に対して，ログデータの層別サンプリングを行うことで解決を図っている．Yin ら [8] は，オフライン推定量の 1 つである MIS 推定量に簡単な修正を加えることで，オフライン推定量の分散が大きくなりすぎるという問題を改善した．また，実ビジネスにおけるモデル形態に着眼し文脈付きバンディットのもとで報酬に一貫性がないケースに対応したモデルも提案されている [9]．これらの研究は，施策実施効果を明らかにしようという立場からの研究といえる．

これに対し，坪井ら [10] は，ユーザごとの施策実施効果の差異に着眼し，機械学習を用いて施策立案する手法を提

案している．この研究は，どのような特徴量を持つユーザ群に対して施策を打つことが有効であるかを推定する手法に関する示唆を与えている．特に効果の解釈性と観察データに含まれるユーザを人為的に選択することによる系統的な誤差（以下，選択バイアス）の除去の両立を考えたものである．効果の解釈性は回帰木のアルゴリズムを用いている因果推論手法 Causal Tree [11] をベースにすることで可能にしている．選択バイアスの除去は，選択バイアスが存在する状況において処置効果が推定可能な手法である Doubly Robust Estimator [12] を用いることで可能にしている．

3.2 アパレル業界における分析研究

アパレル業界においても機械学習は広く活用されるようになっている [13], [14], [15]．ZOZOUSED のデータを機械学習の手法を用いて分析した研究として齊藤らの研究 [16] があげられる．この研究では，季節性アイテムが需要を精確に予測し戦略的に出品される必要性があるという課題に対し，定量的な判断に基づく的確な出品計画を支援するための需要予測手法を提案している．具体的には，シーズン前にワンシーズン全体の販売数予測を行う長期予測モデルと，直近の新たな情報を用いた残差予測による調整を行う短期予測モデルを組み合わせた 2 段階の手法である．

桑田らの研究 [17] では中古ファッション EC における出品価格の決定問題を対象としている．この研究では，出品価格で購入されるか否かを判断する機械学習モデルを構築し，出品価格での購入が予測される商品には価格を徐々に引き上げ，値下げされて購入されると予測された商品には価格を段階的に引き下げることで，適切な出品価格を推定する手法を提案している．具体的には高精度な分類器の 1 つである Random Forests（RF）[18] を用いて 2 値分類器を構築し，その後商品の出品価格の変動による販売結果の予測値を分析し，元の販売結果と異なる結果が得られる出品価格の閾値を見つける．この手法により，値下げを避けつつ，販売結果に影響を与えない範囲内での出品価格の最大値を見積もることができ，実際の商品の出品価格を決定する際の基準となることが期待できる．桑田ら [19] はさらに，中古ファッションアイテムを出品する時点で提示した出品価格の価格保持期間についても検討を行い，機械学習を用いた適正化手法を提案している．中古ファッションアイテムは，出品時点での出品価格で購入されずに売れ残ってしまうケースもあるため，一定期間を経過した場合には徐々に値下げを行うが，これを早く行いすぎると機会損失となり，遅すぎると在庫コストにつながる．桑田らは，この出品価格保持期間について過去の販売履歴データを学習した Natural Gradient Boosting モデルを駆使して適正化する手法を提案し，実データを用いてその効果を検証した．「チェリーピッカー」と呼ばれるセール品ばかりを狙う顧客をどのようにマネジメントするかという問題に対する研究に野際の研究 [20] がある．顧客が一時点で特定の商品を選択する「チェリーピッキング行動」は一般的に店舗やチェーン側にとっては利益をもたらさないが，価格プロモーションは顧客の購買パターンを変えてスイッチやリピートを促進する効果を持つ [21]．そのため，一時的にチェリーピッキング行動をする顧客をそのまま「チェリーピッカー」として特定するのではなく，期間全体の買い物行動を通じてチェリーピッカーかどうかを判断することが重要である．この研究では段階推定による潜在クラス分析を用いてそれを可能にしている．

このように，アパレル業界においても様々な形で機械学習が用いられており，企業の利益向上に寄与している．

3.3 従来研究のまとめと本研究への展開

以上に示したように，ビジネス施策の効果推定に関しては，過去の観察データからいかに正確に効果を推定するかという観点で様々な研究がなされている．これらは，これまでの因果推論の分野が対象としてきた問題設定からのアプローチであり，新たに効果的に実験計画を行い，実験的にデータを取得して分析を行うというアプローチではなかった．また，アパレル業界における機械学習の活用においても，基本的には過去の大量のログデータ（観察データ）を活用して機械学習モデルを構築し，予測や意思決定に活用するというアプローチであり，新たな施策効果の推定を目的とした実験計画を機械学習モデルの分析結果に基づき策定する研究はない．

これに対して本研究では，過去の大量のログデータによって構築した機械学習モデルを駆使して，施策効果の高い顧客グループを特定するための施策実験計画の方法論の提案を行う．具体的には，機械学習モデルが出力する予測値で作られる空間上で Upper Confidence Bound による探索的な最適化を行う手法を提案し，人工データと実ビジネスにおける実証実験を通じ，提案手法の有効性を示した．その結果，担当者の属人的な判断で施策対象ユーザを選定していたこれまでの方法に対し，機械学習モデルを導入することで，施策担当者の経験に過度に依存せずに，かつ比較的シンプルな方法論で施策効果の高いユーザグループを発見することを可能とする施策実験計画が可能であることを示す．

4. 提案手法

4.1 提案への着想

本研究では，大量のログデータ（観察データ）を活用して，高い施策効果が期待できるユーザグループを特定し，評価するための施策実験対象グループを設計する手法を提案する．まず，施策効果を推定するための施策実験における適切な実験計画手法を考える．メジャーな手法としてあ

げられるものがランダム化比較試験である．これはユーザをランダムに施策対象と非対象に分ける手法であるり，ランダム化することで，検証したい効果以外の要因がバランスよく分かれるため，公平に比較することを可能とする．一方，本問題においてこの手法はいくつかの問題点を含む．まず，蓄積された過去の大量の観察データを活用できない点，ユーザ依存の施策効果の推定が行えない点である．これらの問題点を解決するためには，ユーザの過去の購買行動データを活用してクラスタリングを行うことで，似たような特徴を持つユーザ群における施策効果を推定することが可能となる．また，全ユーザを対象に行った場合，コストがかかりすぎるという問題もある．これを解決するためには実験対象をなんらかの方法で絞る必要がある．本研究での目的は施策効果が高いユーザグループの特定であるため，あらかじめ観察データを用いて高い施策効果が期待できるユーザに対象を絞ることが望ましい．

ここで現時点での施策効果を推定するアプローチ手法として統計的因果推論 [22] の枠組みがある．しかし，この手法では施策実施における十分なデータがあることが前提である．本問題で扱う過去の施策実施データは，ある特定の特徴を持ったユーザに対して施策がほとんど実施されておらず，一方である特定の特徴を持ったユーザに施策実施が極端に集中しているという課題がある[*1]．このような状況は，経験的により効果が見込めるユーザに施策を実施しようとした場合に多く発生しうる事象であり，本研究対象の通販サイト以外でも往々にして起こりうる課題である．このように顧客の特徴によって施策対象顧客数が大きく偏る場合，高い精度での施策効果推定が難しい．したがって，他の手段を用いて施策効果の推定を行うことが求められる．つまり，本研究で主に考えるべきことはユーザのクラスタリング方法，過去の施策実施データが不十分な状況下での効果推定手法の 2 点である．まずユーザのクラスタリング方法について，クラスタリング手法には K-means 法 [23] や潜在クラスモデル [24], [25] などの様々な手法があげられる．本研究では実ビジネスで活用するうえでのユーザセグメントの解釈性を考慮し，一般に顧客価値を決める指標である購買単価と利用回数に基づいてセグメント化することを考えた [26]．一方，実際の値には異常値となるものも含まれてしまうため，機械学習モデルを用いて予測購買単価と予測利用回数を算出し，これらを指標にする．機械学習モデルの構築には分析対象の過去の施策に関する情報を含む観察データとは異なる過去のデータを使用する．次に過去の施策実施データが不十分な状況下での効果推定手法について考える．単純に施策対象者数の違いについて考慮せず 2.2 節で述べたように効果を算出した場合，効果の値は，観察データにおいて施策対象者の割合がかなり小さいユーザ群では不確かなものとなってしまう．そこでこのような状況下では効果の分散を考慮して効果の推定を行う必要があるといえる．本研究では分散の考慮には Gaussian Process Upper Confidence Bound（以下，GP-UCB）[27] の考え方を取り入れた．GP-UCB はベイズ最適手法 [28], [29], [30] の 1 つであり，期待値に標準偏差を加味した値を推定値とする手法であり，これにより期待値と不確実性の両方を考慮できる．なお本研究が従来のベイズ最適化手法と異なる点として，ビジネス施策の最適化における実応用上の制約条件下で実施可能な手法であることがあげられる．具体的には，本研究が対象としているビジネス施策の最適化の場合，最適化のための探索空間を適切に定義する必要があり，かつ施策実験の実施コストの観点から一定期間内で同時に実施する施策実験の計画を効果的に組む必要がある[*2]．本研究はこれらの制約条件のもとで 1 つの実施可能な方法論を示しており，この点において従来手法の単純な応用ではないといえる．

4.2 提案手法のプロセス

本研究では，過去の観察データを用いて構築した機械学習モデルによる予測購買単価と予測利用回数の切り口でユーザをグルーピングし施策効果が高そうなユーザ群を定義する．具体的な手順を以下に示す．

[提案手法の手順]

STEP1) 将来の利用ユーザの抽出：

利用の有無を予測する 2 値分類器を用いて，今後の利用があると予測されたユーザのみを抽出する．

STEP2) 予測購買単価と予測利用回数の推定：

STEP1 で抽出されたユーザに対して，類似したユーザ同士でグルーピングを行うために，各ユーザの予測購買単価と予測利用回数を算出し，それに基づいて「予測利用回数が 2 回以上 3 回未満」,「予測購買単価が 3,000 円以上 4,000 円未満」などのように区分を適切に決定する．この区分に従い，ユーザを「予測購買単価」と「予測利用回数」の 2 軸で分類し，グルーピングを行う．

STEP3) 施策対象群の特定：

実験対象となる顧客グループについて，各グループで適切に施策対象群と非対象群に分け，処置群に属するユーザを施策対象とする．

この提案手法の STEP1 において，利用が少なく離反の恐れがあるユーザについては，すでに施策対象となって

[*1] 施策対象ユーザを，データベース上の条件合致検索によって抽出すると，このような偏りが生じやすい．たとえば，「入会○○カ月以内」,「過去○○カ月以内に購買実績がない」などのデータベース上で定義される論理的な検索条件で抽出されているユーザを施策対象とするといったケースである．

[*2] ベイズ最適化は，明示的に与えられたパラメータ空間で，何度も実験と評価を繰り返して最適点を発見していく探索プロセスを想定した手法といえる．これに対し，本研究では，ある期間内でいっせいに実施するビジネス施策の実験に対する計画を対象としており，連続的な探索プロセスというよりは 1 回の実験計画を想定している．

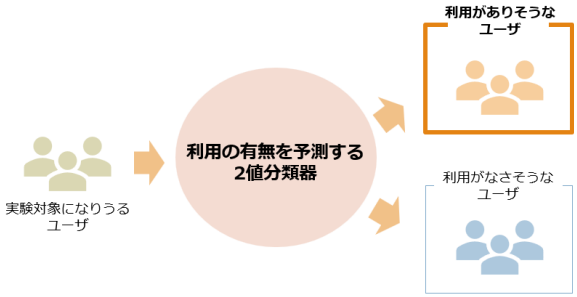


図 2 提案手法 STEP1 のイメージ
Fig. 2 The image of STEP1.

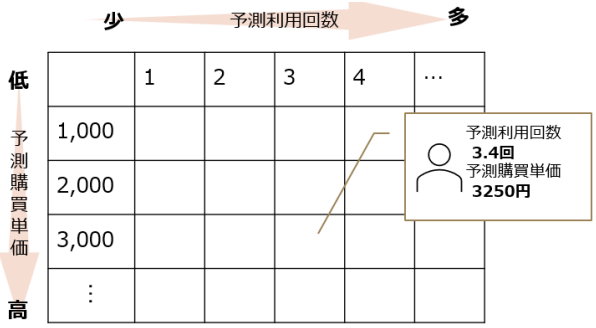


図 3 提案手法 STEP2 のイメージ
Fig. 3 The image of STEP2.

いるため除外する．2 値分類器は，過去のデータを用いて Light gbm [31] により構築する．STEP1 のイメージを**図 2** に示す．

また，STEP2 におけるグルーピングは，ユーザの特徴を表す指標である顧客価値に基づいくユーザグルーピングであるといえる．一般に，顧客価値は購買単価と利用回数によって定められる．これは総購入金額が購買単価と利用回数の積であるからである．一方，実際の値には異常値となるものも含まれてしまうため，予測購買単価と予測利用回数を指標に分類を行う．予測購買単価，予測利用回数においても過去のデータを用いて Light gbm により構築する．STEP2 のイメージを**図 3** に示す．

施策の対象グループには，観察データを基に期待効果が高いと見込まれるグループを定義する．ここで効果とは，施策が打たれたユーザとそうでないユーザ間の施策前後における利用回数の増加分の平均の差分で定義される．一方，効果の値はグループに属する人数や施策有無の比率によって不確かさをともなうため，ここでは効果の期待値にユーザの数や偏りが考慮された標準偏差を加味したものを期待効果とする．効果の期待値を μ，標準偏差を σ，パラメータを α とすると，各グループの Upper Confidence Bound は式 (4) のように定義される．

$$\hat{y} = \mu + \alpha\sigma \tag{4}$$

さらに対象グループの施策が打たれたユーザ群の利用回数の増加分の平均値を μ_A，平均値の分散を σ_A^2，そうではないユーザ群の利用回数の増加分の平均値を μ_B，平均値の分散を σ_B^2 とすると μ と σ は式 (5)，(6) のように定義される．

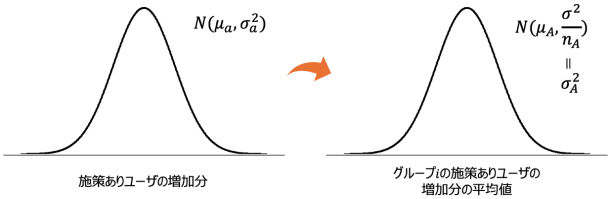


図 4 μ_A，σ_A^2 の導出イメージ
Fig. 4 The derive image of μ_A and σ_A^2.

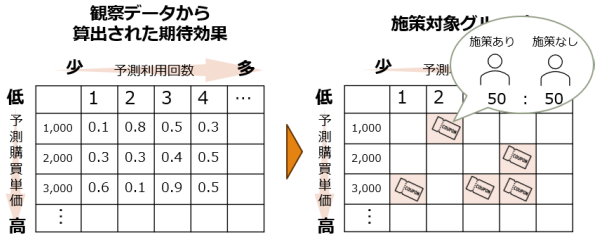


図 5 提案手法 STEP3 のイメージ
Fig. 5 The image of STEP3.

$$\mu = \mu_A - \mu_B \tag{5}$$

$$\sigma = \sqrt{\sigma_A^2 + \sigma_B^2} \tag{6}$$

ここで μ_A，μ_B，σ_A^2，σ_B^2 は利用回数の増加分が正規分布に従っていると仮定して算出する（**図 4**）．つまり，施策が打たれた全ユーザの利用回数の増加分が平均 μ_a 分散 σ_a^2 の正規分布に従っていると仮定したとき μ_A，σ_A^2 は式 (7)，(8) のように算出される．n_A は対象グループのユーザ数である．

$$\mu_A = \mu_a / n_A \tag{7}$$

$$\sigma_A^2 = \sigma_a^2 / n_A \tag{8}$$

このようにして導出した期待効果に対し，値が高い順に施策実験対象者上限人数に収まるようにユーザグループを選択し，各グループ内でランダムに施策対象群と非対象群が 1:1 になるように分ける．STEP3 のイメージを**図 5** に示す．

5. シミュレーション実験

提案手法内で用いるパラメータ α は実データの分散の大きさや値のスケールに依存する．そこで本研究で扱うデータに対し適切にパラメータ α を設定するため，実データに即して生成した人工データを用いてシミュレーション実験を行う．

5.1 実験に用いる人工データセット

本実験ではパラメータ α の決定のみに着目するため，予

測利用回数と予測購買単価を導出する過程は省略し，あらかじめグループが割り振られた状態のデータを生成する．まず各グループに適切なユーザ数を設定し，各ユーザごとに施策対象の有無，施策前の利用回数，施策後の利用回数を設定した．ここでシミュレーション実験において，実データに対し適切にパラメータ α を設定するうえで，人工データにおける効果の期待値と標準偏差が実データと同等の規模感である必要があるため，ユーザ数などの各値は過去の施策実施データを参照して決定した．以下に，具体的な決定方法を説明する．まず施策対象かどうかは，各グループに実データに即した施策実施割合を設定したうえで決定した．まず，ユーザの利用回数はポアソン分布に従うものと仮定する．計数データが従う確率分布としてはより複雑なモデルも考えられるが，本研究で対象とする ZOZOUSED はファッションアイテムを取り扱っており，各ユーザがときどき利用するような EC サイトである．また，このユーザによるサイト利用という事象は，各ユーザが持つ一定確率に従って定常かつ独立に発生すると仮定しても差し支えない[*3]．そのため，定常かつ独立に発生する事象の回数が従う離散確率分布として最も基本的なポアソン分布を仮定し，ユーザの利用回数のサンプリングにはポアソン分布を用いた．ハイパーパラメータ λ は，各グループに実データに即した平均利用回数 a を設定したうえで，$k=1$，$\theta=a$ をハイパーパラメータに持つガンマ分布からサンプリングした．施策後の利用回数は，離反確率 b を設定し，離反の場合は 0，そうでないかつ施策対象でない場合は前述した施策前の利用回数と同様の方法で作成し，施策対象の場合は前述した施策前の利用回数と同様の方法で作成したものに施策効果を決める変数 c をかけたものとした．施策効果 c について，この値は過去の実際の施策効果，あるいは期待値に則り定める必要があるため，ビジネス上の知見により適切に設定する必要がある[*4]．利用回数のサンプリング方法のイメージ図を**図 6** に，人工データの概要を**図 7** に示す．

5.2 パラメータ α の決定

前節で用いた各グループの平均利用回数 a，離反確率 b，施策効果を決める変数 c より真の効果の値 y を定義し，式 (9) に示す．

$$y = a \times (1-b) \times (c-1) \tag{9}$$

[*3] 離散事象のモデルとしては，たとえば「地震の発生」や「電子メールの送受信」など，1 度，事象が発生するとバースト的に連続して事象が発生しやすくなるようなモデルもあるが，本研究で対象とする EC サイトの利用は，このような例には該当しない．一定期間内の事象の発生回数，ある平均回数を持ち，定常かつ独立に生起することを仮定したモデルがポアソン分布であり，ユーザの利用回数が従う確率分布としては最も基本的である．

[*4] 施策効果 c の実際の数字については，ビジネス上の理由から非公開情報となっている．

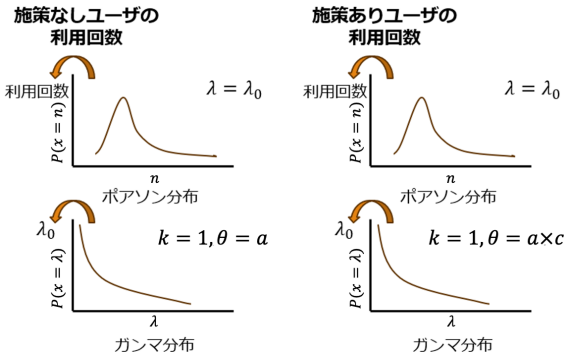


図 6 利用回数のサンプリング方法
Fig. 6 Sampling method for frequency of use.

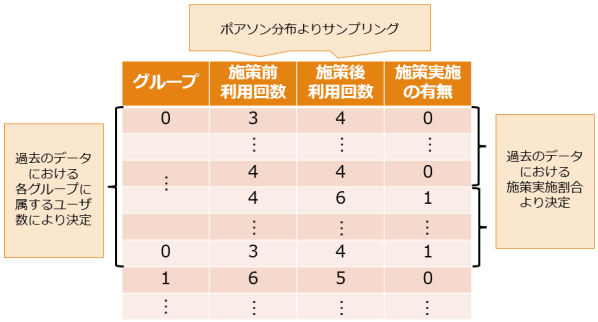


図 7 人工データの概要図
Fig. 7 Overview figure of artificial data.

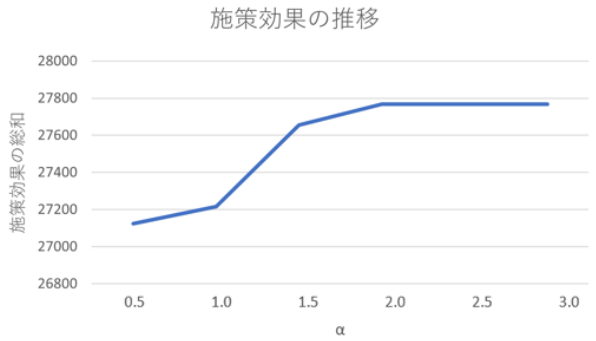


図 8 α と施策効果の総和の関係
Fig. 8 Relationship between α and the sum of measure effects.

施策対象者数の上限を設定し，その範囲内で人工データに提案手法を適用し算出した期待効果が高い順に施策対象グループを決定し，各グループに属するユーザ数に真の効果をかけたものの総和を算出する．この実験を $\alpha = \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$ に対して行い，その結果に基づいて α を決定する．

5.3 実験結果と考察

図 8 に，α を変化させたときの施策効果の総和の推移を示す．図 8 より，α の値が 2.0 以上のときに施策効果の総和が高いことが分かる．一方，本研究における検証実験では，施策利用の上限に応じて施策実施期間を決定する．そのため，効果の不確実性を大きく考慮，つまり探索を重視した実験計画は，施策効果がないかつ施策利用率が高いユー

グループ(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	0	0	**0.636**	0.890	**1.186**	0	**0**
(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
0	0	0	0.636	0.890	1.186	0	0
(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
0	0	0	0.636	0.890	1.186	0	0
(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
0.254	0.381	0.508	1.589	2.225	**2.966**	**3.337**	**3.708**
(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)
0.254	0.381	0.508	1.589	2.225	**2.966**	**3.337**	**3.708**
(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)
0.254	0.381	0.508	1.589	**2.225**	2.966	**3.337**	**3.708**
(49)	(50)	(51)	(52)	(53)	(54)	(55)	(56)
0	0	0	0.636	0.890	1.186	1.335	1.483
(57)	(58)	(59)	(60)	(61)	(62)	(63)	(64)
0	0	0	0.636	0.890	1.186	**1.335**	1.483

図 9 施策対象グループと真の効果
Fig. 9 Target groups for measures and true effects.

ザ群の存在を誘発させ，施策実施期間が短くなるという形で実験に悪影響を及ぼす可能性が高い．したがって，本研究では $\alpha = 2.0$ とする．このとき，シミュレーション実験において選択された施策対象グループは**図 9** の黄色で塗られたグループである．また値は真の施策効果を表しており，効果の値が大きいほど濃い色で色づけられている．ここで効果の値は平均を 1 としたときの倍数を指している．

図 9 より，効果推定値と分散の和が大きい順番にグループを選択した場合，おおむね真の効果が大きいグループが施策対象として選択されていることが分かる[*5]．本研究での対象は「真の施策効果が高いにもかかわらず，過去に施策があまり実施されていない顧客グループがある」という状況において，「顧客グループの施策効果が高い顧客グループ」をいかに発見するかという問題を扱っている．実際に図 9 を見るとグループ (1)，(8) のように真の効果が小さいにもかかわらず，実験対象として選択されているグループが存在している．これは，過去のデータがまだ十分でないために，推定される施策効果の信頼区間が比較的広くなり，標準偏差を大きく考慮していることが原因と考えられる．推定で得られる標準偏差の大小は各グループに所属するユーザ数やそれらの均質性に依存するものであり，実際の適用時にはグルーピング方法が適切であるかどうかに加え，各グループに属するユーザの数を可能な限り確保する必要があることが示唆される．また，グループ (46) のように効果が大きいにもかかわらず選択されていないグループについては，すでに過去に統計的に必要な施策数が得られているためである．すなわちグループ (46) は「施策効果が高い」，かつ「過去の施策実施数が多い」ユーザのセグメントであるいえる．このようなセグメントはすでに効果的な施策対象グループであることが明らかとなっており，新たに追加実験の必要はない．すなわち，このセグメントは新たな試作実施の必要性が少ないが，今回の施策実験の結果，引き続き施策効果が高いユーザグループであるとして施策対象となりうるセグメントである．グループ (46) のような高い施策効果が見込めるグループを抽出することもまた企業経営においては重要な問題であり，今回の施策実験の結果によって効果の低いグループが特定されることで，相対的に施策効果が高いグループであると示されることも実験の重要な成果であると考えられる．また，このグループと類似する特徴量を持つ周辺のグループについては，施策実験の対象として選択されていることも見て取れる．すなわち，実際の適用においてはグルーピングの粒度を大きくすることや，あるいは周辺のグループが施策対象となった場合に同様に選択対象とするなどの工夫も可能であると考えられる．ただし，施策効果の高いユーザセグメントが施策対象グループとなっているか否かは，「施策効果の大きさ」と「過去の施策実施数の多さ/少なさ」のバランスによって決まってくるが，これに加えて「どのくらいの数の施策実験を行うことができるか」という「実験規模」にも依存する．すなわち，より多くのユーザに対して実験的な施策を打つことができれば，真の効果が高いユーザグループを実験候補に加えることが可能となるが，反面，施策効果が低いグループへの施策実験も増えるというトレードオフの問題が生じるため，実際場面では施策実施コストを考慮した検討が必要である[*6]．

6. 検証実験

本章では，ZOZOUSED における過去の施策実施データを基に施策実験対象者を決定し，実際に施策を実施し効果検証を行った結果について示す．

6.1 実験条件

提案手法内で構築する利用有無の 2 値分類器，購買単価予測モデル，利用回数予測モデルは，2020 年に購買を行ったユーザ群を対象とし学習を行った．説明変数には RFM（Recency，Frequency，Monetary）[32] を主として，**表 1** に示した変数を用いた．パラメータは人工データを用いたシミュレーション実験より $\alpha = 2$ と設定した．ここで表 1 内の USED は ZOZOUSED を，TOWN は ZOZOTOWN を意味する．観察データには 2021 年に購買を行ったユーザを対象とし，施策は 2021 年 12 月に実施したものを用

[*5] 本実験で適用している，探索的な実験で用いられる Upper Confidence Bound は，推定量の平均と標準偏差から得られる信頼上限である．具体的には，過去に得られているユーザ数と施策実施結果のデータを用いて計算された信頼上限であり，点推定値としての平均値と区間推定としての標準偏差を個別に評価することは難しい．本研究では「施策効果が高い顧客グループ」を発見するための手法の提案を目的としており，「施策効果の大きさ」を示す推定量の平均と「過去の施策実施数の多さ/少なさ」に影響される標準偏差の双方を用いる Upper Confidence Bound を用いることが問題設定に照らして適切であるといえる．

[*6] なお，各ユーザセグメントに属するユーザ数や施策対象者数の実際の数字については，これらの数字が施策実施の規模などの推測を可能とする理由で企業秘密となっており，非公開情報のため，掲載することができない．

表 1 説明変数一覧
Table 1 Explanatory variables list.

説明変数	概要
F_TIMES_USED	2020 年に USED で購買を行った回数
M_USED	2020 年の USED での総購買金額
F_ITEMS_USED	2020 年の USED での購買商品数
DIFF0_USED	USED で 2020 年内最後に購入を行った日から 2020 年 12 月 31 日までの日数
DIFF1_USED	USED で 2020 年内最後に購入を行った日とその前の購入日との間隔
F_TIMES_TOWN	2020 年に TOWN で購買を行った回数
M_TOWN	2020 年の TOWN での総購買金額
F_ITEMS_TOWN	2020 年の TOWN での購買商品数
DIFF0_TOWN	TOWN で 2020 年内最後に購入を行った日から 2020 年 12 月 31 日までの日数
DIFF1_TOWN	TOWN で 2020 年内最後に購入を行った日とその前の購入日との間隔

いた．

本実験では直近 1 年以内に購買を行ったユーザを対象に提案手法を適用し，2023 年 5 月にポイント付与施策を実施し検証を行う．また，現状施策対象になりにくいユーザの施策効果を知ることが目的のため，現状対象に含まれている新規ユーザは実験対象からは除外する．なお，純粋な施策効果を検証するため，直近で他の施策が打たれたユーザも除外する．施策実施期間は施策利用の上限に応じて決定する．また，利用有無の 2 値分類器，購買単価予測モデル，利用回数予測モデルが，それぞれ次の分析に用いるために十分な精度が得られているか否かについては，企業担当者に確認を依頼し，問題がないことを確認した[*7]．

6.2 実験結果と考察

6.2.1 定量的分析

前節に基づいて実際に施策を実施し，現段階における効果を検証した．現段階では施策実施後 6 カ月間の購買履歴のみが得られるため，施策前 1 年間における LTV の 2 分の 1 との差分を施策前後における増加分として効果を算出した．**表 2** に各グループのユーザ数と施策効果の総和，施策利用率を示す[*8]．なお，各グループの分割には lightGBM による施策後 6 カ月の予測利用回数，予測購買単価を用いている．予測購買単価を用いる理由として，顧客の嗜好性や優良顧客度を大まかに測る指標として有用であることがある．ZOZOUSED では 2024 年 10 月末時点で 7,500 以上のブランドのファッションアイテムの中古品を取り扱っているため，商品によって価格帯が大きくばらつくことが考えられる．具体的には，シャツやブラウスなどのトップス

[*7] なお，これらのモデルの精度そのものについては，説明変数からどの程度詳細に予測が可能であるかを直接的に示してしまうことから，共同研究先企業側での非公開情報となっており記載することができない．

[*8] なお，各グループに属している具体的な顧客数については企業情報で非公表であるため，施策利用者分布の記述にとどめている．

表 2 検証結果
Table 2 Verification result.

グループ	ユーザ数	施策利用者分布	総効果指数
(1)	0.092	0.017	−19.968
(2)	0.915	0.082	2.072
(3)	0.692	0.048	25.872
(4)	0.577	0.043	91.560
(5)	0.152	0.041	2.924
(6)	0.109	0.024	−5.27
(7)	2.442	0.144	−9.681
(8)	1.135	0.038	−68.801
(9)	1.404	0.089	19.305
(10)	0.473	0.031	−43.416
(11)	0.941	0.017	12.792
(12)	1.225	0.048	59.684
(13)	0.816	0.019	−22.176
(14)	1.487	0.206	−15.156
(15)	2.539	0.153	107.850
合計			137.591

系とコートやジャケットなどのアウター類など種類によって価格帯が異なる商品や，ブランド品のシューズやバッグ，アクセサリーなどのブランド品も扱っており，それらのどのような商品を購入するユーザであるかによって購買単価は異なっている．実際に ZOZOUSED では，それらのユーザ顧客の嗜好の差異を細かい購入品目の類似性に基づく顧客クラスタリングを用いて分析する試みも行っているが，中古商品が基本的に 1 品物であり，まったく同じ商品が複数存在しないという点で，通常の「同じ商品を購入している顧客は，嗜好が類似している」という顧客クラスタリングや協調フィルタリングで用いられる考え方を用いた手法の適用が難しい．そこでユーザのグルーピング手法として，顧客の嗜好性や優良顧客が反映されていると考えられる「購買単価」を用いることとした．次に区分化粒度の設定について，各グループに含まれるユーザ数が同程度であり，かつ試作効果を推定するために十分なユーザ数が含まれるように決定した．すなわち，機械学習によるクラスタリング手法とは異なり，「施策実験の対象者を決めるためのグループ単位」を定めるために実験者が決定する必要があるといえる．ここで各グループの推定精度のみを考慮した場合，施策効果が「予測購買単価」と「予測利用回数」によって急激に変化するならば区分化粒度を小さくし，一方で緩やかに変化するならば区分化粒度を大きくしてもよい，ということができる．しかし細かく区分化した場合にはそれぞれのグループにおいて施策効果推定が可能となる施策実験の対象者数を増やす必要があるが，クーポン発券などのコストは限られているため多すぎるグループ数への区分化は現実的ではない．仮に施策効果である利用回数の増加分の変動が「予測購買単価」と「予測利用回数」に対して比較的緩やかであるという仮定が成り立つ場合，大きな

区分粒度であってもビジネス視点では問題ないグルーピングが可能であるといえる．以上より，本研究では各グループにある程度のユーザ数が含まれるよう，統計学的なデータ数の観点から決定した[*9]．また，表 2 のユーザ数比率は，各グループに所属するユーザ数の全体平均値を 1.00 とした場合の相対的なユーザ数の比率を示している．また，施策利用者分布は各施策利用者に対する各グループの利用率を，総効果指標は各グループでの増加利用回数の合計を定数倍したものを表す[*10]．

表 2 より，効果のあるユーザ群の存在が確認できた．現状施策が打たれていないユーザの中から，施策効果があると判明したことは企業にとって有益な情報であるといえる．効果指数の値にはばらつきがあるが，今後より長期的なスパンで改めて検証することでさらに効果が安定して増大することが見込める．また，施策効果指数の総和の合計が正の値であったことから，本実験における全体の施策効果はプラスに働いたといえる．一方で，観察データに基づいて算出した期待効果とは異なる値をとった．この要因として 2 つ考えられる．1 つは年による需要の変動や季節要因がもたらすものである．もう 1 つは短期的な効果検証に基づいている点である．古着は生活必需品ではないため，購入頻度は比較的少ない傾向にある．それゆえに短期的な効果検証による結果は，同一グループであってもユーザ間のばらつきが大きく効果の値は安定しないと考えられる．

また，施策利用者分布について，これらはおおよそユーザ数に依存していることが分かる．対象が多ければ多いほど利用者数は増加するので当然ではある一方，グループ (14) はユーザ数に対して施策利用率が高く，グループ (11), (13) はユーザ数に対して施策利用率が低いことがうかがえる．グループ (14) は効果が負であることから，チェリーピッカーを多く含む可能性が考えられる．グループ (11), (13) について，ユーザ数に対して施策利用率が低いということは，すなわち処置群の中に含まれる施策利用者がかなり少ないことを意味するため，処置群と対照群との比較が施策効果を表すとはいい難い．

6.2.2 定性的分析

前項では各グループごとに効果を算出し分析を行った．本項では実験対象グループをセグメント化の軸として用いた予測利用回数と予測購買単価に基づいてグループを割り振り定性的な分析を行う．予測利用回数を 5 段階，予測購買単価を 3 段階に区切って各グループを割り振ると**表 3** のようになる．なお，予測購買単価，予測利用回数ともに，段階の数が大きくなるほど単価ならびに回数が増加することを示している．また，表に示すそれぞれの数値はあくまで分割した段階を示しており，回数や単価そのものを示す値ではない．

この割振りを基に施策効果について定性的な考察を行う．表 3 を基に各マスの平均効果を算出し**表 4** に，ユーザ数の和を**表 5** に，施策利用率の和を**表 6** に示す．なお，平均効果を算出する際は，各グループの平均ユーザ数あたりの施策効果指数を算出したうえで，マスに属するグループの平均を平均効果とした．

表 4 より，予測利用回数が異なるユーザ間で大きく施策効果指標が異なることが明らかになった．これは，必ず継続的に購買が発生するわけではないファッション通販サイトにおいて，どの程度の頻度でサイトを利用するかはユー

表 3 予測利用回数と予測購買単価に基づいたグループ割り振り
Table 3 Group assignment based on predicted frequency of use and predicted purchase price per unit.

		予測利用回数				
		1	2	3	4	5
予測購買単価	1		(6), (14)	(5)	(1)	
	2		(15)	(7), (9)		(2)
	3	(11)	(12), (13)	(8)	(4)	(3), (10)

表 4 予測利用回数と予測購買単価に基づいた平均施策効果指数
Table 4 Average measure effect based on predicted frequency of use and predicted purchase price per unit.

		予測利用回数				
		1	2	3	4	5
予測購買単価	1		−29.448	19.254	−217.459	
	2		42.273	9.061		2.265
	3	13.591	10.760	−60.594	158.564	−271.824

表 5 予測利用回数と予測購買単価に基づいたグループのユーザ数
Table 5 Number of users in a group based on predicted frequency of use and predicted purchase price per unit.

		予測利用回数				
		1	2	3	4	5
予測購買単価	1		1.596	0.152	0.092	
	2		2.539	3.846		0.915
	3	0.941	2.041	1.135	0.577	1.165

表 6 予測利用回数と予測購買単価に基づいたグループにおける施策利用者分布
Table 6 Measure utilization for the group based on predicted frequency of use and predicted purchase price per unit.

		予測利用回数				
		1	2	3	4	5
予測購買単価	1		0.230	0.041	0.017	
	2		0.153	0.233		0.082
	3	0.017	0.067	0.038	0.043	0.079

[*9] なお，具体的な分割基準についてはこれらの数字が共同研究先企業における販売状況の推測を可能とする理由で企業秘密となっており，非公開情報のため，掲載することができない．
[*10] 各ユーザセグメントに属するユーザ数そのものについては，ビジネス規模や施策実施数などの直接的な値を推測可能な数字でありため，企業側で非公開情報となっている．そのため，本研究ではユーザ数ではなく，その比率によって相対的なユーザ分布によって，各セグメント規模の傾向を示すこととした．

ザのファッションに対する嗜好の強さに強く影響しており，予測利用回数が 1 段階異なっただけでも，そのグループに属するユーザ特性がかなり変化するためと考えられる．たとえば，利用回数が 5 回程度のユーザは 2，3 カ月に 1 度の頻度でファッションアイテムを購入するユーザであり，ボーナス時期や季節の変わり目といった季節的なイベントを待たずに購入をしてくれるユーザである．この点で，利用回数が 3 回程度のユーザとは顧客の優良度や特性がかなり異なると考えられる．実際，予測購買単価が 3，利用回数が 5 のグループでは効果指標がマイナスとなっているが，これはこのグループに属するユーザが特に強いファッションへの関心を持っており，かつ ZOZOUSED への愛着が強いため施策が実施されなくても購買を行うようなユーザであることが示唆される．一方で利用回数が 1 段階下がることで，ファッションへの強い関心はあるものの他のサイトでの購入も行うユーザが多く存在することが想定され，そのようなユーザに対しては施策を実施することで ZOZOUSED の利用を促進することができたと推察される．

7. 考察

7.1 ユーザセグメント手法の妥当性

本研究では，ユーザのセグメンテーションに際し，各ユーザの予測購買単価と予測利用回数を算出し，それに基づいて「予測利用回数が 2 回以上 3 回未満」，「予測購買単価が 3,000 円以上 4,000 円未満」などのようにグルーピングを行った．これは，一般に顧客価値は購買単価と利用回数によって定められる一方，実際の値には異常値となるものも含まれてしまうためである．ここで，似たような顧客価値どうしでのセグメンテーションを行うという本来の目的を達成するための手段として予測購買単価と予測利用回数の 2 軸でのセグメンテーションが妥当かどうかの検証が必要である．そこで本節では，予測利用回数と予測購買単価が実際の値に対してどのように分布しているかを見ることで各軸の妥当性を検証する．

図 10，**図 11**，**図 12** に予測購買単価に対する実際の購買単価の分布を 3 段階に分けて，**図 13**，**図 14**，**図 15**，**図 16**，**図 17** に予測利用回数に対する実際の利用回数の分布を 5 段階に分けて示す．横軸は実際の値，縦軸は度数を表しているが，両者ともは企業情報となるため表記しない．また，図 10〜12，図 13〜17 ではそれぞれ横軸と縦軸のスケールは同じである．

図 10〜17 より，両者とも値が大きくなるとともに分布の山が右にずれていっていることが分かる．したがって，これらの軸は顧客セグメンテーションにおいて適切に機能しているといえる．

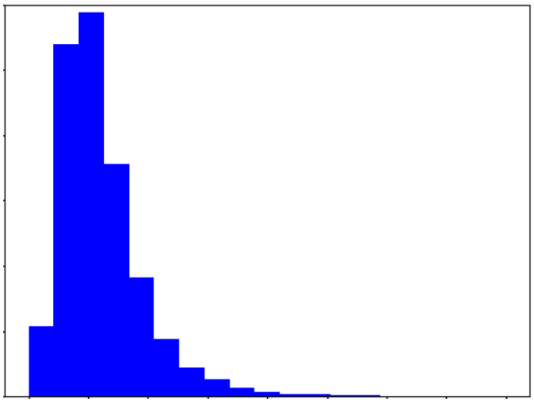


図 10 予測購買単価レベル 1 の実際購買単価の分布
Fig. 10 Distribution of actual purchase price per unit of predicted purchase price level 1.

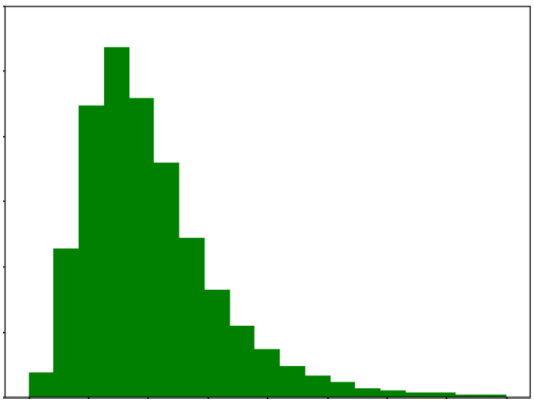


図 11 予測購買単価レベル 2 の実際購買単価の分布
Fig. 11 Distribution of actual purchase price per unit of predicted purchase price level 2.

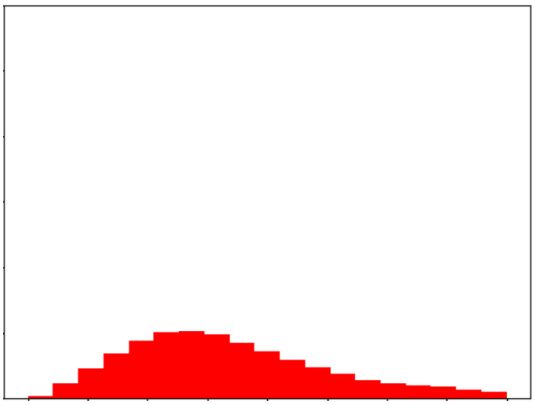


図 12 予測購買単価レベル 3 の実際購買単価の分布
Fig. 12 Distribution of actual purchase price per unit of predicted purchase price level 3.

7.2 提案手法の有効性

検証実験より，提案手法によって未開拓の顧客群の中から施策効果がある顧客群を抽出できることが明らかになった．まず本研究手法で得られる分析結果の信頼性について議論する．仮に機械学習で学習された「予測購買単価」ならびに「予測利用回数」の回帰モデルの精度が不十分で

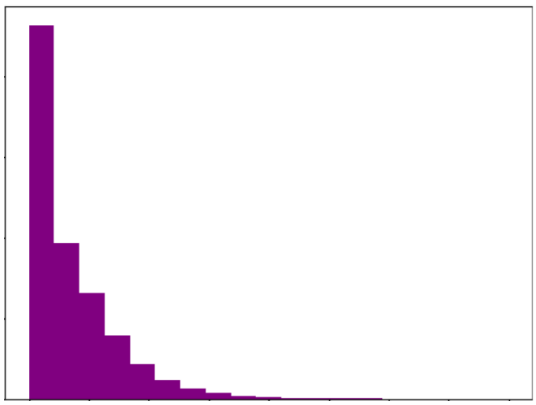


図 13 予測利用回数レベル 1 の実際利用回数の分布
Fig. 13 Distribution of actual use frequency for predicted use frequency level 1.

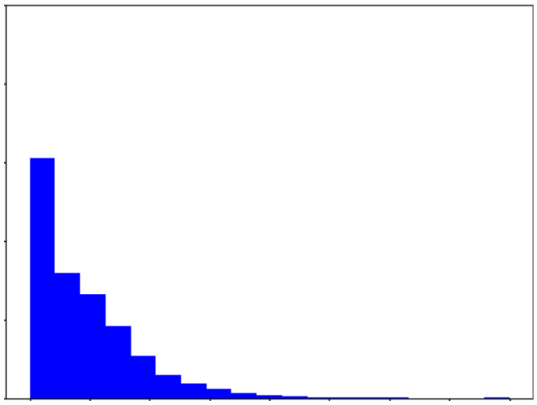


図 14 予測利用回数レベル 2 の実際利用回数の分布
Fig. 14 Distribution of actual use frequency for predicted use frequency level 2.

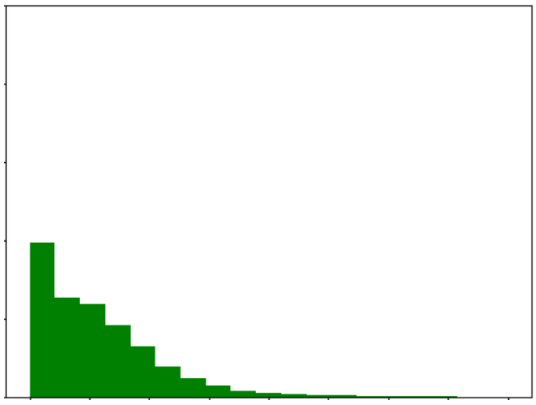


図 15 予測利用回数レベル 3 の実際利用回数の分布
Fig. 15 Distribution of actual use frequency for predicted use frequency level 3.

あった場合，得られるグルーピング結果は完全にランダムなものとなり，グループ間の施策効果に差異は生じないものと考えられる．しかし実証実験ではグループごとに異なる施策効果が得られており，「予測購買単価」および「予測利用回数」の推定精度はある程度担保されていたこと，ならびにこれらの指標が顧客の嗜好や傾向を表す指標として

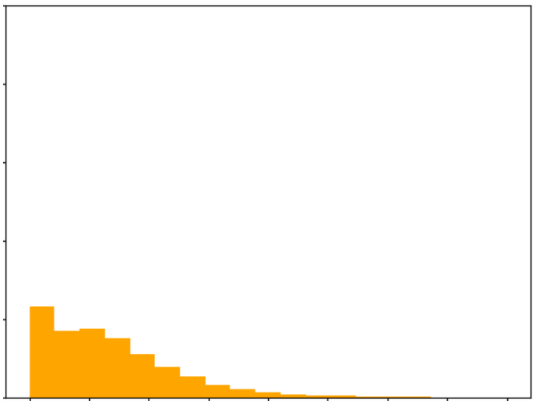


図 16 予測利用回数レベル 4 の実際利用回数の分布
Fig. 16 Distribution of actual use frequency for predicted use frequency level 4.

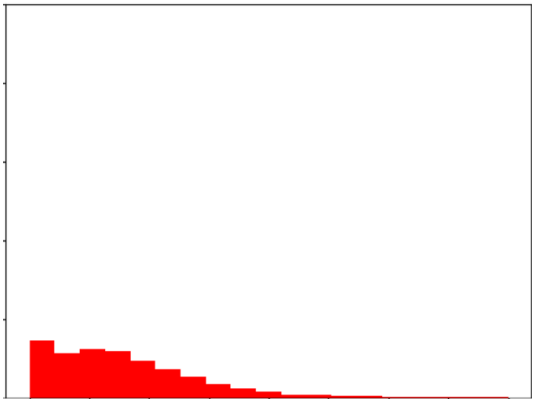


図 17 予測利用回数レベル 5 の実際利用回数の分布
Fig. 17 Distribution of actual use frequency for predicted use frequency level 5.

有効であったことが示唆される．加えて，これらの指標の予測精度が大幅に改善された場合には実験計画の効率性が向上し，さらに有意義な示唆を得られることが期待される．また，マーケティングにおいて顧客開拓は重要である [33]．顧客を顧客ロイヤルティ別に 5 段階で分けた場合，多くの新規ユーザや離反予備軍に対する施策は 1 段階目の顧客を 2 段階目にするためのものである．しかし 1 段階目のみに着目した開拓が長期的かつ全体的に見たときに，必ずしも最適な開拓方法とはいえない．本提案手法ではその他の段階における顧客開拓に着目し，その中に効果のある顧客群を特定することを可能にしたものであり，これは企業に大きな発見を与えたといえる．

8. 結論と今後の課題

本研究では，ファッション通販サイト ZOZOTOWN 内で古着を販売する ZOZOUSED を対象として，ポイント付与施策未実施のユーザ群の中から高効果が望めるユーザ群を特定するための定量的指標に基づいた実験計画手法を提案した．提案手法では，予測購買単価と予測利用回数の切り口でユーザをセグメント化し，各セグメントにおける施

策効果の平均と分散を基に施策対象とするセグメントを決定した．これにより，どのような施策実施データに対しても効率的に高効果ユーザ群を特定することが可能になった．

人工データセットを用いたシミュレーション実験において，提案手法において適切なハイパーパラメータ設定を行い，またその設定のうえで提案手法が施策実施実験において施策効果が高いユーザの探索を行いつつも的確に高効果ユーザ群を特定できることを示した．さらに，検証実験により，提案手法を用いることによって実際に高効果ユーザ群を特定することができることが明らかとなった．

本研究の成果により，施策によって顧客ロイヤルティが高いユーザ群の中に，さらなる向上が望めるユーザ群を新たに発見することが可能になった．これにより，新規ユーザの定着だけでなく顧客ロイヤルティ向上における施策効果を持つユーザ群の特徴抽出が可能になり，実務上の貢献も期待できる．

今後の課題として，施策効果の推定における観察データと実験データの融合方法の考案があげられる．本手法によって実験データを得ることができた．このデータと過去の観察データをを融合することによってより多くの情報を持つ新たな観察データを構築することができ，また次の施策実施実験への活用が望める．

謝辞 本研究の一部は，JSPS 科学研究費 21H04600, 24H00370 の助成を受けたものです．

参考文献

[1] 平野洋介, 楊 添翔, 雲居玄道, 阿部 永, 立花徹也, 後藤正幸：顧客成長を促す施策立案のための特徴転移型クラスタリングモデル，情報処理学会論文誌，Vol.62, No.10, pp.1704–1715 (2021).

[2] ZOZOUSED：ブランド古着の通販･買取- ZOZOTOWN，入手先 ⟨https://zozo.jp/zozoused/⟩

[3] Hoekstra, J.C. and Huizingh, E.K.: The lifetime value concept in customer-based marketing, *Journal of Market-Focused Management*, Vol.3, pp.257–274 (1999).

[4] Pearl, J.: *Causal inference in statistics: An overview* (2009).

[5] Deaton, A. and Cartwright, N.: Understanding and misunderstanding randomized controlled trials, *Social Science & Medicine*, Vol.210, pp.2–21 (2018).

[6] Narita, Y., Yasui, S. and Yata, K.: Efficient Counterfactual Learning from Bandit Feedback, *Yale: Cowles Foundation Working Papers* (2018) (online), available from ⟨https://api.semanticscholar.org/CorpusID:52179881⟩.

[7] Kallus, N., Saito, Y. and Uehara, M.: Optimal Off-Policy Evaluation from Multiple Logging Policies, *ArXiv*, Vol.abs/2010.11002 (2020) (online), available from ⟨https://api.semanticscholar.org/CorpusID:224814085⟩.

[8] Yin, M. and Wang, Y.-X.: Asymptotically Efficient Off-Policy Evaluation for Tabular Reinforcement Learning, *International Conference on Artificial Intelligence and Statistics* (2020) (online), available from ⟨https://api.semanticscholar.org/CorpusID:210942713⟩.

[9] Wang, Y.-X., Agarwal, A. and Dudík, M.: Optimal and Adaptive Off-policy Evaluation in Contextual Bandits, *Proc. 34th International Conference on Machine Learning*, Precup, D. and Teh, Y.W. (Eds.), *Proc. Machine Learning Research*, Vol.70, pp.3589–3597, PMLR (2017) (online), available from ⟨https://proceedings.mlr.press/v70/wang17a.html⟩.

[10] 坪井優樹，阪井優太，鈴木佐俊，後藤正幸：Causal Tree に基づく選択バイアスを考慮した頑健な条件付き平均処置効果推定手法の提案，情報処理学会論文誌，Vol.64, No.9, pp.1399–1412 (2023).

[11] Wager, S. and Athey, S.: Estimation and inference of heterogeneous treatment effects using random forests, *Journal of the American Statistical Association*, Vol.113, No.523, pp.1228–1242 (2018).

[12] Funk, M.J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M.A. and Davidian, M.: Doubly robust estimation of causal effects, *American Journal of Epidemiology*, Vol.173, No.7, pp.761–767 (2011).

[13] Xiao, H., Rasul, K. and Vollgraf, R.: Fashion-mnist: A novel image dataset for benchmarking machine learning algorithms, arXiv preprint arXiv:1708.07747 (2017).

[14] Luce, L.: *Artificial intelligence for fashion: How AI is revolutionizing the fashion industry*, Apress (2018).

[15] Xia, M., Zhang, Y., Weng, L. and Ye, X.: Fashion retailing forecasting based on extreme learning machine with adaptive metrics of inputs, *Knowledge-Based Systems*, Vol.36, pp.253–259 (2012).

[16] 齊藤芙佑，山下 遥，佐々木北都，後藤正幸：2 段階の機械学習予測モデルに基づく季節性のある中古アパレル商品の需要予測に関する一考察，情報処理学会論文誌，Vol.63, No.11, pp.1684–1696 (2022).

[17] 桑田 和，杉崎智哉，三川健太，後藤正幸：販売履歴データに基づく中古ファッションアイテムの出品価格推定モデルの提案，情報処理学会論文誌，Vol.62, No.2, pp.796–808 (2021).

[18] Breiman, L.: Random forests, *Machine learning*, Vol.45, pp.5–32 (2001).

[19] 桑田 和，三川健太，佐々木北都，後藤正幸：機械学習アプローチに基づく中古ファッションアイテムの価格保持期間適正化モデルの提案と実証的効果検証，情報処理学会論文誌，Vol.63, No.1, pp.218–230 (2022).

[20] 野際大介：潜在推移分析を用いたチェリーピッカーの特定と理解：ファッション EC サイトへの段階推定法の応用，行動計量学，Vol.46, No.2, pp.73–85 (2019).

[21] 杉田善弘，上田隆穂，守口剛ほか：プライシング・サイエンス：価格の不思議を探る．

[22] 岩崎 学：統計的因果推論 (統計解析スタンダード)，朝倉書店 (2015).

[23] Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R. and Wu, A.Y.: An efficient k-means clustering algorithm: Analysis and implementation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.24, No.7, pp.881–892 (2002).

[24] McCutcheon, A.L.: *Latent class analysis*, Vol.64, Sage (1987).

[25] Collins, L.M. and Lanza, S.T.: *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*, Vol.718, John Wiley & Sons (2009).

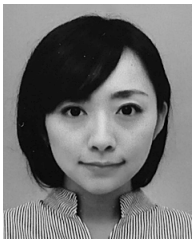
[26] Luarn, P. and Lin, H.-H.: A customer loyalty model for e-service context., *J. Electron. Commer. Res.*, Vol.4, No.4, pp.156–167 (2003).

[27] Srinivas, N., Krause, A., Kakade, S.M. and Seeger, M.: Gaussian process optimization in the bandit set-

ting: No regret and experimental design, arXiv preprint arXiv:0912.3995 (2009).
[28] Snoek, J., Larochelle, H. and Adams, R.P.: Practical bayesian optimization of machine learning algorithms, *Advances in Neural Information Processing Systems*, Vol.25 (2012).
[29] Dorie, V., Hill, J., Shalit, U., Scott, M. and Cervone, D.: Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition (2018).
[30] Brodersen, K.H., Gallusser, F., Koehler, J., Remy, N. and Scott, S.L.: Inferring causal impact using Bayesian structural time-series models, *The Annals of Applied Statistics*, Vol.9, No.1, pp.247–274 (online), DOI: 10.1214/14-AOAS788 (2015).
[31] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.-Y.: Lightgbm: A highly efficient gradient boosting decision tree, *Advances in Neural Information Processing Systems*, Vol.30 (2017).
[32] 阿部 誠：RFM データを用いた顧客生涯価値の算出—既存顧客の維持介入と新規顧客の獲得，マーケティングジャーナル，Vol.34, No.1, pp.73–90 (2014).
[33] Ang, L. and Buttle, F.: Managing For Successful Customer Acquisition: An Exploration, *Journal of Marketing Management*, Vol.22, No.3-4, pp.295–317 (online), DOI: 10.1362/026725706776861217 (2006).

中村 友香

1999 年生．2022 年早稲田大学創造理工学部経営システム工学科卒業．現在，同大学大学院創造理工学研究科経営システム工学専攻修士課程在学中．機械学習を用いたデータ分析に関する研究とその実応用に興味を持つ．



山極 綾子 （学生会員）

1994 年生．2016 年早稲田大学創造理工学部経営システム工学科卒業．2021 年同大学大学院修士課程修了．現在，同大学院創造理工学研究科経営デザイン専攻博士後期課程在学中．機械学習を用いたデータ分析に関する研究とその実応用に興味を持つ．

佐々木 北都 （正会員）

1984 年生．2010 年青山学院大学経営学部二部経営学科卒業．2012 年同大学大学院国際マネジメント研究科国際マネジメント専攻修士課程修了．経営管理修士（MBA）．2017 年株式会社クラウンジュエル（現，ZOZO）入社．データサイエンティストとして主にアパレルリユース事業のビジネス課題の解決，および機械学習プロダクトの開発に従事．

後藤 正幸 （正会員）

1969 年生．1994 年武蔵工業大学大学院修士課程修了．2000 年早稲田大学大学院理工学研究科博士課程修了．博士（工学）．1997 年早稲田大学理工学部助手．2000 年東京大学大学院工学研究科助手．2002 年武蔵工業大学環境情報学部助教授．2008 年早稲田大学創造理工学部経営システム工学科准教授．2011 年同大学教授．情報数理応用とデータサイエンス，ならびにビジネスアナリティクスの研究に従事．著書に，『入門パターン認識と機械学習』コロナ社 (2014)，『ビジネス統計—統計基礎とエクセル分析』オデッセイコミュニケーションズ (2015) 等．IEEE，INFORMS，電子情報通信学会，人工知能学会，日本経営工学会，経営情報学会等各会員．