

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

桑田 和^{1,a)} 三川 健太^{2,b)} 佐々木 北都^{3,c)} 後藤 正幸^{1,d)}

受付日 2020年12月3日, 採録日 2021年10月8日

概要: 本研究で対象とする中古ファッション EC サイトでは、売れ残りのリスクを避けるために出品から一定期間が経過したアイテムの値下げを行う仕組みを導入しており、初期価格の保持期間の設定が1つの課題である。従来、値下げは利益への影響が強いものの経験的な方法によって実施されており、過去の販売履歴データの分析に基づいた、初期価格保持期間の設定の適正化が期待されている。また、多様性の高いファッションアイテムをつねに同じタイミングで値下げすることは合理的ではない。一方、初期価格保持期間は対象アイテムが誰かに購入されるまでの期間（掲載期間）を考慮する必要がある。これは購入する顧客が現れた際に確定するため、掲載期間を事前に正確に予測することは一般に難しい。そこで本研究では、アイテムの特徴量から掲載期間の予測分布を推定する機械学習モデルを構築することで、適切な初期価格保持期間を決定する方法を提案する。提案手法では、アイテムごとに掲載期間における販売確率を推定後、アイテムの出品月を変化させて推定を繰り返すことで初期価格保持期間を決定する。本研究では、当該サイトの過去の販売履歴データに提案手法を適用し、その予測モデルとしての有効性を示すと同時に、初期価格保持期間を決定する。そして、その期間を用いて実際にアイテムを出品する実証実験を行い、得られた販売履歴データの分析により実運用における提案手法の有用性を検証する。

キーワード: 中古ファッション EC サイト, 出品価格保持期間, Natural Gradient Boosting, 販売履歴データ, 値下げ

A Proposing Decision Model and Empirical Effect Verification of Sales Period with List Prices for Second-hand Fashion Items Based on Machine Learning Approach

IZUMI KUWATA^{1,a)} KENTA MIKAWA^{2,b)} HOKUTO SASAKI^{3,c)} MASAYUKI GOTO^{1,d)}

Received: December 3, 2020, Accepted: October 8, 2021

Abstract: The second-hand fashion EC site focused in this paper marks down the prices of items that have passed a certain period since they were exhibited. Since the markdown of price has a strong effect on profits, it is important to set a proper period of exhibiting each item with the list price. In addition, it is not rational to mark down every price of items at the same timing. On the other hand, the expected sales period of items should be considered for setting proper period. It is, however, difficult to predict it accurately on the point estimation manner. Therefore, we propose a method to decide a proper period by estimating the probability distribution of the sales period. In the proposed method, after estimating the probability on the sales period for each item, the proper period is decided by changing the listing month of the item and repeating the estimation. We apply the proposed method to the actual sales history data, show the effectiveness of the proposed method, and determine the proper periods. Finally, we experiment with selling real items using the decided periods, and consider the usefulness of our method in actual operation with obtained sales history data.

Keywords: second-hand fashion EC site, sales period on list price, Natural Gradient Boosting, sales history data, markdown

1. はじめに

近年、多くの業界でECサイトが普及し、多くの消費者がECサイトで購買を行っている。それにとともに、ECサイトを運営する企業では膨大な販売履歴データが蓄積され、これらのデータを活用した顧客の行動に関する分析 [1], [2], [3] や、アイテムの特徴に関する分析 [4] などの研究がさかに行われている。また、ECサイトが扱う商品（以下、アイテム）は新品だけではなく、中古品も多く扱われている。中古アイテムを扱う多くのECサイトでは、出品するアイテムの売れ残りを防ぐため、値下げを実施する場合がある。収益性の観点からは、より多くの売上を確保するため、可能な限り値下げをせずに販売することが望ましい。しかし、在庫コストや出品アイテムの鮮度といった観点からは、いつまでもアイテムを保持することはできないため、可能な限り早く販売することも必要となる。そのため、ファッションアイテムの需要を把握することは重要であり、販売予測の観点からの研究も行われている [5]。一方で、出品価格を維持する期間の適切な設定も重要な課題となっている。ここで、出品価格とは、アイテムを出品した際の初期価格を指す。一般にファッション業界において、初期価格の状態は「プロパー」と呼ばれており、出品価格の保持期間を「プロパー期間」と定義する。その適正化を具体的に議論するために本研究では、ファッション通販サイト ZOZOTOWN 内で古着を販売する ZOZOUSED におけるプロパー期間の決定問題を対象事例とする。ZOZOUSED では、ユーザから古着を買い取り、中古品として再販売する事業を展開している。また、出品から一定期間経過したアイテムの価格を段階的に下げる仕組みを採用しており、プロパー期間の決定は非常に重要な課題の1つである。ZOZOUSED におけるプロパー期間や出品価格に関するイメージを図 1 に示す。

一般に、ファッションアイテムは、ブランドや季節性、カテゴリといった性質を有する。そのため、アイテムの多様性に合わせたプロパー期間の決定が必要である。一方、アイテムの購入はユーザの意思決定によるものであり、各アイテムの掲載期間（出品から誰かに購入されるまでの期間）にはサイトを訪れたユーザの判断が反映されている。すなわち、掲載期間はそのアイテムの購入されやすさ（人気度合い）を表した値であると考えられる。そのため、プ

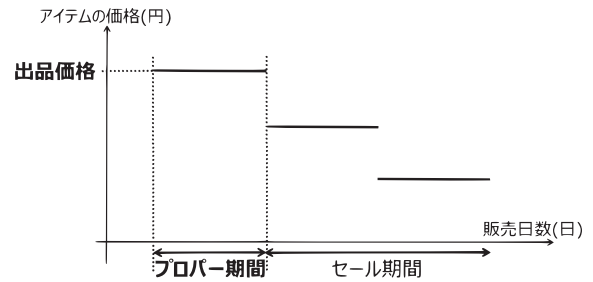


図 1 プロパー期間のイメージ

Fig. 1 The image of proper period.

ロパー期間の決定には売れやすさを意味する掲載期間を考慮する必要があるが、以下の課題が存在する。1つ目は、アイテムの掲載期間はばらつきが非常に大きいため正確な1点予測が困難であるという点である。2つ目は、季節性のあるアイテムも存在するために、全アイテムに対しつねに同じプロパー期間を適用することは合理的ではないという点である。これらの解決のため、前者に対しては確率分布を用いて掲載期間のばらつき方を予測することを考え、後者に対してはアイテムの出品月ごとにプロパー期間を設定し、適正な価格保持期間の決定を考える。

上記の議論のもと、本研究では、各出品月におけるアイテムの適切なプロパー期間の決定方法を提案する。提案手法ではまず、アイテムの特徴量と出品月を入力として各アイテムの掲載期間の予測分布を推定するモデルを構築する。予測モデルには、勾配ブースティングの1つである Natural Gradient Boosting (NGBoost) [6] を用いる。さらに、予測モデルの入力のうち、出品月のみを変化させて予測を繰り返すことで、月ごとに適切なプロパー期間を決定する。本研究では、提案手法の有効性を検証するため、ZOZOUSED の販売履歴データに提案手法を適用し、分析結果に対する考察を行う。加えて、提案手法の実運用における有効性を検証するため、新規出品アイテムに対して決定したプロパー期間を適用し、販売結果に対する分析と考察を行う。

2. 準備

2.1 関連研究

ECサイトを対象とした研究では、顧客に関する行動分析などが多くなされている。山田ら [7] は、顧客の購買履歴データとアンケートデータに基づいて、共分散構造分析を用いて季節ごとに適した商品の売り方を示している。野際ら [8] は、ECサイトにおける購買履歴データを活用し、潜在推移分析を用いてセール品ばかりを狙う顧客（チェリーピッカー）を特定する枠組みを提案している。一方、杉山ら [9] は、商品価格が最も高い（販売開始）時期に商品を購入する顧客（イノベーター）の検出モデルを提案している。

また、ファッション業界では、アイテムの売上予測や需要予測を目的とした研究が多く、予測精度や計算速度といった観

¹ 早稲田大学
Waseda University, Shinjuku, Tokyo 169–8555, Japan
² 湘南工科大学
Shonan Institute of Technology, Fujisawa, Kanagawa 251–8511, Japan
³ 株式会社 ZOZO
ZOZO, Inc., Chiba 261–7116, Japan
a) ikdsk14243150@toki.waseda.jp
b) mikawa@info.shonan-it.ac.jp
c) hokuto.sasaki@zozo.com
d) masagoto@waseda.jp

点から様々な手法が用いられている [5]. Thomassey ら [10] は、クラスタリング手法と決定木を用いて、販売履歴データに基づく中期的な売上予測について研究している. Ni ら [11] は、トレンド傾向、季節性といった要因を考慮した販売履歴データと外部要因に基づく長期売上予測と短期売上予測を組み合わせた2段階予測を行っている. 一方, Au ら [12] は、モデルの計算コストの観点から研究を行い、遺伝的アルゴリズムを用いたニューラルネットワークによる短期的な売上予測を行っている. さらに, Choi ら [13] による SARIMA とウェーブレット変換を組み合わせた手法による販売時系列データに基づく売上予測についての研究も行われている.

このように売上予測に関する研究が多くある一方、各アイテムの販売価格を予測するモデルとして仁ノ平ら [14] の研究がある. 仁ノ平らは、設定した出品価格のもとでアイテムを販売した場合の最終的な販売価格を予測するモデルを構築し、実販売データに対する予測精度の面からその有効性を示している.

さらに、アイテムの特徴分析モデル [4] や、EC サイトを効率的に閲覧するためのコーディネート支援システム [15]、顧客の購買履歴データとアンケートデータに基づくブランドの推薦システム [16] などが提案されている.

以上のように、ファッション業界では顧客の購買行動分析や、アイテムの需要や価格の予測に関する研究が多くなされている. しかし、機械学習を用いた値下げの仕組み、特にプロパー期間の決定に焦点を当てた研究は少なく、議論の余地があるといえる.

2.2 Natural Gradient Boosting

Natural Gradient Boosting (NGBoost) は勾配ブースティング [17] に自然勾配法 [18] を用いた手法である. 自然勾配法は、通常の勾配法に比べてパラメータの収束速度が大幅に速くなることが示されている [19].

NGBoost は、「弱学習器」, 「任意の次元数のパラメータ θ を持つパラメトリックな確率分布」, 「スコア関数」の3つの構成要素を持ち、弱学習器を用いて確率分布のパラメータを学習し、推定された確率分布と観測値を用いてスコア関数を計算する. そして、スコア関数を最適化するようにして自然勾配を学習する.

いま、 $l-1$ 個目の弱学習器の学習で得られた条件付確率分布のパラメータを $\theta^{(l-1)}$ 、予測されたパラメータによる確率分布と観測値 v から算出されるスコア関数を $S(\theta^{(l-1)}, v)$ とする. このとき、 l 個目の弱学習器において、自然勾配 $g^{(l)}$ は式 (1) を用いて学習される. ただし、 \mathcal{I}_S はリーマン計量であり、スコア関数 S によって変化する. また、 $\nabla_{\theta} S$ は θ における S の通常の勾配である.

$$g^{(l)} = \mathcal{I}_S(\theta^{(l-1)})^{-1} \nabla_{\theta} S(\theta^{(l-1)}, v) \quad (1)$$

NGBoost では「予測の不確かさ」を扱うために、確率分

布 (予測分布) の形で予測を行う. 通常、1点予測値を出力する一般的な回帰モデルは、正規ノイズを仮定した場合の期待値を1点で予測することが多い. これに対し、目的変数の予測分布を求めるアプローチはベイズ統計においてよく行われてきたが、ベイズ推定は比較的シンプルなモデル以外には適用が難しいという特徴がある. NGBoost では、出力 (目的変数) の予測分布のパラメータが、入力 (特徴ベクトル) の関数で与えられる構造で表現し、弱学習器を用いて推定された (個々の入力に対する) パラメータに基づいて出力の予測分布を推定する. 1点予測値を得たい場合には、推定された予測分布の期待値を予測値とすることもできるが、その場合にも、その1点予測値に対し確率値を信頼性尺度として付与することができる.

3. 対象問題

3.1 問題設定

本研究では、古着事業を展開している ZOZOUSSED における値下げの仕組みを対象事例とする. 当該サイトでは、アイテムの売れ残りを防ぐために、出品から一定期間 (プロパー期間) を経過したアイテムの価格を段階的に下げる仕組みを採用している. 前述のとおり、収益性の観点からは値下げをせずに購入されることが望ましいが、管理コストの観点からは、できるだけ早く販売することが望ましいため、プロパー期間の検討は重要であると考えられる. 一方、ZOZOUSSED における従来のプロパー期間は経験的な方法によって決定されている. そのため、これまで値下げされてから販売されたアイテムは、プロパー期間をより長く設定していれば値下げされずに (出品価格のまま) 販売できた可能性がある. 反対に、プロパー期間を長くしても、単に掲載期間が延びるだけで、出品価格では売れなかった可能性もある. そのため、各アイテムに対し適切なプロパー期間を設定することが望まれている.

ZOZOUSSED では多数 (9,000 以上) のブランドの、幅広い (常時 60 万点以上の) アイテムを扱っているため、アイテムによって適切なプロパー期間が異なると考えられる. また、同じアイテムでも出品時期や使用状態によって販売傾向が異なるため、つねに同じプロパー期間で値下げをすることは合理的ではない. したがって、アイテムの多様性を考慮したプロパー期間の決定が必要である.

加えて、前述のように、プロパー期間の決定には、アイテムの売れやすさが反映された掲載期間を考慮する必要がある. ここで、販売履歴データにおける掲載期間に着目し、類似アイテムの比較や単純な分析を行うことで適切なプロパー期間を決定することを考える. しかしながら、以下の理由により、単純な分析を行うだけでは適切なプロパー期間の決定は難しいと考えられる.

- 出品条件が同一のアイテムが少ないため、出品月のみが異なる同一アイテム間の比較や使用状態のみが異な

るアイテム間の比較といった、同一条件下での正確な比較は困難である。

- 掲載期間に寄与する特徴量はアイテムによって異なるため、アイテムごとにそのような特徴量を発見し、それによってプロパー期間を決定することは非効率である。

そこで過去の販売履歴データを用いた機械学習により、掲載期間の予測モデルを構築することで適切なプロパー期間を決定することを考える。しかし、さらに以下の2つの課題が考えられる。

- 各アイテムの掲載期間は多様な値をとりうるために、回帰系のモデルを用いてこれらを高精度で1点予測することは困難である。
- 季節性のあるアイテムも存在するため、すべてのアイテムに対してつねに同じプロパー期間で値下げをすることは適切でない。

そのため、この2つの課題を解決したプロパー期間の決定方法が必要である。前者の問題点に対しては、掲載期間の予測分布を推定することを考える。また、後者の問題点に対しては、出品月を入力に加えた予測モデルの構築により、出品月ごとにプロパー期間を設定することを考える。適切なプロパー期間の決定により、従来の仕組みによる値下げと比較して、出品価格からの値下がり率（以下、オフ率）が変化する可能性があり、アイテムの値下がり回避や売上・粗利額の増加が期待できる。

3.2 事前分析

本節では、本研究で対象とするZOZOUSEDの販売履歴データについて、ブランドに関する分析、出品月に関する分析、掲載期間に関する分析を行い、より詳細なアイテムの特徴を把握する。1件のデータにはあるアイテムが買い取られてからユーザに購入されるまでの情報が与えられている。ここでは2019年1月1日から2019年12月31日に販売された3,295,297件を分析対象とした。

まず、ブランドに関する分析を、各アイテムに対して付与されている「ブランドクラス」を用いて行う。ブランドクラスとは、アイテム管理のために現場で導入されている管理単位であり、ブランドクラス1の平均販売単価が高く、ブランドクラス8の平均販売単価が低くなるように分類されたアイテムの8つのクラスである*1。ただし、ブランドクラスはアイテムのブランドと使用状態に基づいて分類されているため、同じブランドであってもアイテムによって異なるクラスに分類されることもある。また、平均販売単価とは前年の販売履歴データを用いて計算しており、アイテムのブランドクラスは出品される前に付与されている。

ブランドクラスは販売価格というアイテムの重要な特徴量に基づいて分類されるため、各アイテムの特徴を良く反

表 1 ブランドクラスごとの掲載期間と販売価格の中央値
Table 1 Median of sales periods and sales price for each brand classes.

ブランドクラス	アイテム割合 [%]	掲載期間	販売価格
1	1.96	0.74	2.81
2	1.03	1.67	8.33
3	3.34	1.74	4.51
4	8.83	1.26	2.99
5	29.03	1.15	1.70
6	45.53	1.00	1.00
7	2.77	1.22	0.68
8	7.50	0.52	1.74
計	100.00	-	-

映した分類ができていると期待できる。ここで、販売価格とはアイテムが購入された際の価格を指す。そこで、以下では各クラスに所属するアイテムの販売傾向の違いを把握するため、ブランドクラスごとの掲載期間と販売価格について分析を行う。

表 1 にブランドクラスごとの掲載期間と販売価格の中央値を示す。ただし、各値はアイテムの所属割合が最も多いブランドクラス6における中央値を1.00とし、それに対する比率を示す*2。表 1 より、ZOZOUSEDで扱われるアイテムの多くはブランドクラス5と6に分類されていることが分かる。また、ブランドクラス2と3の掲載期間の中央値は他のブランドクラスに比べて、大幅に長いことが分かる。これと表 1 より、ブランドクラス2, 3は全体と比べてアイテム数が比較的少なく、購入者も現れにくいクラスであることが考えられる。さらに、ブランドクラス2~7における販売価格の中央値を比較すると、クラス2の方が高く、クラス7の方が低いことが分かる。以上のことから、掲載期間と販売価格を用いて比較を行うと、各ブランドクラスはそれぞれ異なった販売結果になるといえる。

次に、出品月に関する分析を行う。ZOZOUSEDでは、半袖のアイテムやセーターといった季節性のあるアイテムも扱うため、アイテムの季節性に合わせて出品することもある。そのため、アイテムに付与されている特徴量のうち、出品月と季節ラベルはアイテムの季節性を反映した特徴量であることが考えられる。特に出品月は季節ラベルに比べて粒度が細かく、より詳細な分析が可能であると考え、本節では出品月を対象とした分析を行う。具体的には、ブランドクラスと同様、販売傾向の違いを把握するため、出品月ごとの掲載期間と販売価格を分析する。表 2 に出品月ごとの掲載期間の中央値、販売価格の中央値を示す。ただし、各値はアイテムの所属割合が最も多い10月における中央値を1.00とし、それに対する比率を示す*3。

表 2 より、アイテムは6月、10月、11月に多く出品され

*1 各ブランドクラスの平均販売単価の実測値は非公開情報であるため、記載することができない。

*2 実際の値は非公開情報であるため、比率による記載をした。

*3 実際の値は非公開情報であるため、比率による記載をした。

表 2 出品月ごとの掲載期間と販売価格の中央値

Table 2 Median of sales periods and sales price for each listed month.

出品月	アイテム割合 [%]	掲載期間	販売価格
1	6.47	0.93	1.28
2	7.32	1.07	1.25
3	6.43	0.76	1.20
4	7.19	0.83	1.20
5	9.01	0.97	0.85
6	9.78	1.21	0.90
7	8.18	1.14	0.90
8	8.29	1.14	0.95
9	8.92	1.07	0.99
10	10.09	1.00	1.00
11	9.81	0.79	0.99
12	8.52	0.72	1.05
計	100.00	-	-

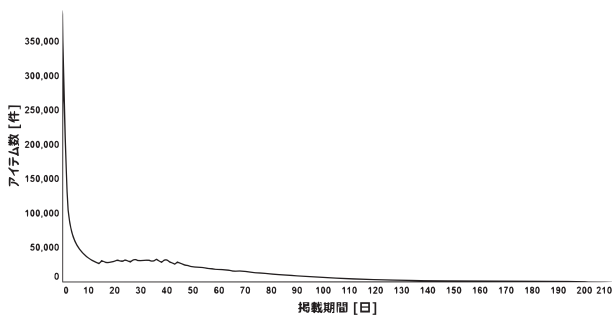


図 2 掲載期間の分布

Fig. 2 The distribution of sales period.

ていることが分かる。また、6月、7月、8月に出品されたアイテムの掲載期間の中央値は他の出品月と比べて、アイテム全体の掲載期間より比較的長いことが分かる。一方、11月と12月に出品されたアイテムの掲載期間の中央値はアイテム全体と比べて比較的短いことが分かる。このことから、出品月によって掲載期間が変化することが考えられる。さらに、販売価格の中央値を比較すると、1月、2月は比較的高価格であるのに対し、5月、6月、7月は比較的低価格で販売されている。これは、冬服の相場が比較的高く、夏服の相場が比較的安いことが原因であると推測できる。したがって、各アイテムの販売価格は出品月によっても変化すると考えられる。以上のことから、掲載期間と販売価格を用いて比較を行うと、アイテムは出品月によってそれぞれ異なった販売結果になるといえる。

次に掲載期間に関する分析を行う。本研究では、掲載期間を考慮したプロパー期間の決定を行うため、その特徴を把握する必要がある。そこで、図 2 に全アイテムの掲載期間の分布を示す。ただし、分布の特徴をとらえやすくするため、アイテム数が少ない 200 日以降は除いて分析を行った。図 2 より、掲載期間は 0 日が最も多く、出品初日に購入されていることが読みとれる。また、その概形は指数分

布に類似していることが分かる。

以上のことから、ブランドクラスと出品月による分類はアイテムの季節性を考慮しつつ、販売結果における特徴をとらえているといえる。そこで、本研究では、ブランドクラスと出品月を用いて、アイテムごとの適切なプロパー期間の決定を目指す。

4. 提案手法

4.1 着想

前述のとおり、ZOZOUSED で扱われるアイテムは多様性が高いため、掲載期間に寄与する特徴量はアイテムによって異なることが考えられる。そのため、多くのアイテムの掲載期間に寄与する特徴量を個々に把握することは膨大な時間を要し、効率的ではない。そこで、類似したアイテムごとにその特徴量を把握するため、アイテムをグルーピングすることを考える。本研究では、アイテム管理のために現場で導入されている「ブランドクラス」を用いてアイテムをグルーピングする。ブランドクラスはアイテムの大きな特徴である価格帯を中心として類似したアイテムを分類した特徴量であるため、アイテムの特徴をうまくとらえられると考えられる。これにより、各アイテムの掲載期間に寄与する特徴量を把握しやすくなると期待できる。

さらに、3.2 節で述べたとおり、各アイテムの掲載期間は多様な値をとりうるため、掲載期間そのものを正確に予測することは困難である。加えて、掲載期間の分布は指数分布に類似していることが明らかとなっている。そのため提案手法では、各アイテムの掲載期間に指数分布を仮定し、2 章で述べた NGBoost を用いて掲載期間を予測する。これにより、区間予測が可能となり、掲載期間を確率分布の形で予測できるため、この分布をプロパー期間の決定に活用できると考えられる。

一方、前述のように、各アイテムは季節性や使用状態などによって販売傾向が変化することが考えられる。そのため、全アイテムをつねに同じプロパー期間で値下げをすることは適切ではない。また、3.2 節より、アイテムの販売結果は出品月により異なることが明らかである。すなわち、適切な掲載期間は出品月で変化すると考えられるため、出品月がアイテムの販売結果を大きく左右する特徴量であると考えられる。そこで、出品月を掲載期間に強く寄与する特徴量と見なし、各アイテムの出品月のみを変化させたときの販売確率の変動を比較・分析することで適切なプロパー期間の決定を考える。

4.2 概要

提案手法ではまず、ブランドクラスを用いてアイテムのグルーピングを行う。ブランドクラスごとに、アイテムの特徴量と出品月を入力とした NGBoost を学習し、各アイテムを出品した際の掲載期間の予測分布を推測する。この

とき、予測に用いる分布として指数分布を仮定する。

次に、プロパー期間を決定したいアイテムに対しては、出品月のみを変化させ、構築した予測モデルを適用して予測分布を推定する。推定されたアイテムの掲載期間の予測分布から、予測分布の α %点の掲載期間（以下、設定掲載期間）を算出する。ただし、 α は値下げをせずに（出品価格のまま）販売したいアイテムの割合（目標値）として、実務的観点から決定される値である。そして、ブランドクラスごとに得られた各アイテムの設定掲載期間（ α %点）の平均値を算出し、適切なプロパー期間とする。固定する出品月を1月から12月まで変化させることで、各ブランドクラスにおける出品月ごとのプロパー期間を決定する。

4.3 提案手法のアルゴリズム

いま、予測モデルの構築に用いる N 個の学習用アイテムと、プロパー期間の決定に用いる M 個のアイテムについて考える。また、出品月を予測モデルの入力（特徴ベクトル）に加えることにより、出品月別の掲載期間の予測分布が算出できるモデルを構築する。出品月は1月から12月までとし、1-hot ベクトルの形で予測モデルへの入力とした。ただし、各ブランドクラスの学習データには全出品月（1月から12月）に出品したアイテムが十分に存在するものとする。

N 個のアイテムのうち、 n 番目のアイテムの特徴量と出品月をベクトル化したものを $\mathbf{x}_n \in \mathbb{R}^D$ 、掲載期間を $y_n \in \mathbb{R}^+$ とする。ここで、 D は予測モデルの入力次元数とする。ブランドクラスは前述のとおり8クラスとし、 M 個のアイテムのうち、ブランドクラス i のアイテム数を M_i ($i = 1, \dots, 8$) とする。すなわち、 $\sum_i M_i = M$ である。また、出品月は1月から12月までとする。このとき、詳細なアルゴリズムは以下のとおりである。

Step1) 掲載期間の予測分布を推定するモデルの構築

各アイテムの特徴量と出品月のベクトル \mathbf{x}_n を説明変数、掲載期間 y_n を目的変数として、ブランドクラスごとにNGBoostを適用し、各アイテムの予測分布を推定するモデルを構築する。

Step2) プロパー期間の決定

Step2-1) ブランドクラスを $i = 1$ とする。

Step2-2) ブランドクラスが i であるアイテムを抽出し、出品月を $j = 1$ とする。

Step2-3) 抽出された M_i ($M_i \leq M$) 個のアイテムに、Step1で構築したブランドクラス i のモデルを適用する。

Step2-4) 各アイテム m に対して、適用したモデルから推定された出品月 j の予測分布の下位 α %点の掲載期間 s_m^{ij} を求め、アイテム m の設定掲載期間（ α %点）とする ($1 \leq m \leq M_i$)。

Step2-5) M_i 個の s_m^{ij} の平均値を算出し、ブランドクラス i のプロパー期間 t_{ij} とする。

Step2-6) $j = 12$ のとき、Step2-7へ。 $j < 12$ であれば、

$j = j + 1$ として Step2-3 へ戻る。

Step2-7) $i = 8$ のとき終了する。 $i < 8$ のとき、 $i = i + 1$ として Step2-2 へ戻る。

5. 実データを用いた提案手法の評価

本章では、ZOZOUSEDにおいて過去に販売されたアイテムの販売履歴データに提案手法を適用して、他の予測モデルや従来のアイテムのグルーピング方法と比較することで、提案手法の有用性について評価を行う。また、提案手法を販売履歴データに適用し、適切なプロパー期間を決定する。

5.1 データ概要

本分析では、ZOZOUSEDで2018年1月1日から2019年11月17日に出品された301,129件のアイテムの販売履歴データを用いる。モデル構築に用いる学習データは2018年に出品されたアイテム（72,007件）、プロパー期間の決定に用いるテストデータは2019年に出品されたアイテム（229,122件）とした。ただし本分析では、プロパー期間の決定によって値下げされずに販売される可能性の高い、オフ率（出品価格からの値下がり率）0%と10%で販売されたアイテムを分析対象とした。また、アイテムは買取・販売実績が2回以上あるアイテムを対象とした。

掲載期間の予測モデルの構築においては、出品価格、定価、買取価格、男性向け商品/女性・子供向け商品、使用状態、使用日数などのアイテムの特徴量に出品月を加えた特徴量 ($D = 28$) を用いた。

5.2 評価1：テストデータに対する予測精度の比較による評価

本節では、提案手法において掲載期間の予測モデルとして用いるNGBoostの有用性を検証するために、予測精度の観点から他の予測モデルとの定量的な比較を行う。

5.2.1 分析条件

予測モデルの比較手法にはNGBoostと同じ勾配ブースティング手法の1つである、LightGBM [20] を用いる。この手法は、決定木のアルゴリズムを利用しており、1点予測が可能なモデルである。そのため、区間予測はできないが、構造化データに対しては非常に有効な予測モデルとして広く知られている。そこで、NGBoostで推定した予測分布の期待値を1点予測値 \hat{y}_t と見なし、評価指標は比較手法に合わせて式(2)に示す平均平方二乗誤差 (RMSE) を用いる。ただし、テストデータ数を T 、実観測値を y_t 、予測値を \hat{y}_t とする。

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2} \quad (2)$$

5.2.2 分析結果

予測モデルによるRMSEを表3に示す。

表 3 RMSE による予測モデルの評価

Table 3 Evaluation of predicting model by RMSE.

ブランド クラス	学習データ		テストデータ	
	NGBoost	LightGBM	NGBoost	LightGBM
1	22.73	23.77	26.76	27.53
2	29.01	30.40	29.12	30.46
3	23.38	23.98	24.57	25.75
4	21.14	21.50	13.85	14.74
5	16.03	16.03	12.94	13.98
6	15.92	15.95	12.87	13.82
7	10.20	12.44	14.43	14.39
8	14.59	15.20	12.16	12.91

(太字は RMSE が小さい方を示す)

表 3 より、ほとんどのクラスにおいて、学習データ、テストデータともに提案手法で用いた NGBoost の方が予測精度が高いことが分かる。これにより、掲載期間の予測においては、モデル構築の際に確率分布を推定する NGBoost を適用することが有用であると考えられる。

5.3 評価 2：テストデータに対する予測分布の比較による評価

5.2 節では NGBoost の有用性を示すため、構造化データに対しては高い予測精度を持つことが知られている LightGBM との比較を行った。しかしながら、分布を推定し予測を行う NGBoost を適用することの優位性を示すためには、1 点予測を行う LightGBM との比較だけではなく、分布を推定する手法との比較も必要であると考えられる。そこで、指数分布のパラメータを最尤推定した場合の掲載期間の予測結果と NGBoost を用いた掲載期間の予測結果を比較することで、NGBoost の優位性を示す。

5.3.1 分析条件

提案手法と同様に、比較手法もブランドクラスごとにモデルを作成することとし、ブランドクラスごとに 1 つの指数分布を仮定してパラメータを最尤推定する。予測分布を用いて評価を行うため、評価指標には式 (3) に示す対数損失を用いる。ただし、テストデータ数を T 、掲載期間の観測値を y_t 、入力ベクトル (アイテムの特徴量 + 出品月) を \mathbf{x}_t とし、 $\hat{P}(y_t|\mathbf{x}_t)$ は \mathbf{x}_t のもとでの y_t の予測確率である。

$$L = \frac{1}{T} \sum_{t=1}^T \{-\log \hat{P}(y_t|\mathbf{x}_t)\} \quad (3)$$

5.3.2 分析結果

NGBoost と最尤推定を用いた場合の対数損失を表 4 に示す。表 4 より、以下の 2 点において NGBoost の方が有用であるといえる。

- ほとんどのブランドクラスにおいて、NGBoost の対数損失の方が小さい。
- 最尤推定法を用いると、同じブランドクラスに属する

表 4 対数損失による予測モデルの評価

Table 4 Evaluation of predicting model by Log-loss.

ブランド クラス	学習データ		テストデータ	
	NGBoost	最尤推定	NGBoost	最尤推定
1	2.87	3.49	4.20	4.32
2	4.04	3.97	5.46	4.17
3	3.24	3.73	3.21	3.87
4	2.96	3.53	3.04	3.37
5	2.62	3.26	2.72	3.26
6	2.54	3.21	2.64	3.17
7	3.64	2.93	2.49	3.15
8	2.18	2.93	2.49	3.15

(太字は対数損失が小さい方を示す)

アイテムの掲載期間はすべて同じになってしまう。

前者については、表 4 に示すとおり、評価指標として設定した対数損失が、学習データに対してはブランドクラス 2 と 7 を除いて、テストデータに対してはブランドクラス 2 を除いて NGBoost の方が小さい。このことから、ほとんどのアイテムにおいて、NGBoost の方が掲載期間の予測分布の推定精度が高いことが示された。

後者について、NGBoost ではアイテムごとの予測分布を推定するため、設定掲載期間 (α %点) もアイテムによって異なる。しかしながら、最尤推定でブランドクラスにつき 1 つの分布のパラメータを推定すると、このブランドクラスのアイテムの設定掲載期間はすべて同じ結果となる。本提案手法は実運用を目的としており、ビジネスサイドでの結果の納得感も重要な観点である。本研究で対象とした中古ファッションアイテムは多様性が高く、同じブランドクラスでも季節性や価格によって掲載期間は多様な値をとる。そのため、同じブランドクラスのアイテムの掲載期間がすべて同じ予測結果ということは、ビジネスサイドでの納得感が薄いと考えられる。この点でも、NGBoost は有用であるといえる。

5.4 評価 3：予測精度に基づくブランドクラスの有用性の評価

本節では、提案手法においてアイテムのグルーピングを行う際に用いるブランドクラスの有用性を検証するために、ZOZOUSSED で現在用いられているグルーピング方法を比較対象とし、予測精度の観点から定量的な比較を行う。

5.4.1 分析条件

ZOZOUSSED では現行のシステムにおいて、提案手法と同様にアイテムのグルーピングを行い、グループごとにプロパー期間を決定している。そこで、現行のグルーピング方法*4を比較手法とする。NGBoost により推定された掲載期間の予測分布を用いて評価を行うため、評価指標には式 (3) に示した対数損失を用いる。

*4 詳細は非公開情報であるため、予測精度の結果のみを示す。

表 5 各グルーピング方法による対数損失

Table 5 Log-loss for each grouping method.

グルーピング方法	対数損失
ブランドクラス	3.75
現行のルール	3.89

表 6 ブランドクラスごとの α [%]

Table 6 The rate α for brand classes [%].

ブランドクラス							
1	2	3	4	5	6	7	8
68	59	49	45	46	43	35	61

5.4.2 分析結果

提案手法における、ブランドクラスによるグルーピング方法と ZOZOUSED で採用されている現行のグルーピング方法を用いて、2章で示した NGBoost を構築した際のテストデータに対する対数損失を表 5 に示す。

表 5 より、ブランドクラスごとに予測モデルを構築することで高い予測精度を得られることが分かる。このことから、ユーザの意思決定が反映されている掲載期間は、アイテムの価格帯を考慮したグループごとに予測することが有用であると考えられる。

評価 1 から 3 より、ブランドクラスを用いてアイテムをグルーピングし、NGBoost を適用してモデルを構築する提案手法は、アイテムの掲載期間の予測に有用であるといえる。

5.5 評価 4：プロパー期間の決定

最後に、提案手法を用いてアイテムの適切なプロパー期間の決定を行う。

5.5.1 分析条件

提案手法において、アイテムのグルーピング方法にはブランドクラスを用い、掲載期間の予測モデルには NGBoost を用いた。NGBoost の弱学習器の数は事前分析より 500 と設定した。また、NGBoost の学習と予測に用いる分布には指数分布を仮定した。

さらに、プロパー期間の決定の際にあらかじめ設定する α の値は、現場担当者との協議により「プロパー期間内に販売したいアイテムの割合」として、現行のプロパー期間内での販売実績に基づいてブランドクラスごとに設定した。設定した割合 α を表 6 に示す。

5.5.2 分析結果

構築した NGBoost を用いて、出品月を 1 月から 12 月に変化させて予測を繰り返すことで決定したプロパー期間を表 7 に示す。

表 7 より、ブランドクラス 1 や 2 はプロパー期間が比較的長く、ブランドクラス 7 は比較的短いことが分かる。前述したとおり、ブランドクラスは主に平均販売単価が高い

表 7 決定したプロパー期間 [日]

Table 7 The decided proper period [days].

出品月	ブランドクラス							
	1	2	3	4	5	6	7	8
1 月	39	58	34	22	19	18	14	24
2 月	39	63	42	29	32	32	14	26
3 月	39	71	37	23	19	18	14	25
4 月	39	64	32	22	19	18	14	25
5 月	39	57	30	22	19	16	12	24
6 月	38	59	34	22	19	18	14	24
7 月	38	59	34	22	19	18	14	24
8 月	41	58	34	22	19	18	14	24
9 月	41	57	34	22	19	18	14	24
10 月	38	58	41	24	22	19	16	24
11 月	39	58	35	22	21	20	14	24
12 月	39	58	34	22	19	18	14	24

順にグルーピングされている。ZOZOUSED では、価格帯が高いアイテムほど、長くサイトに掲載することでサイトの魅力度が向上すると考えている。そのため、この結果はサイトの魅力度の観点からも適切なものとなっている。

以上のことから、実用的なプロパー期間決定モデルが構築できたといえる。

6. 提案手法の実運用による検証実験

本章では、提案手法の実運用における有効性を示すため、ZOZOUSED で新たに出品されるアイテムに対して、前章で決定したプロパー期間を実際に適用した検証実験の結果を示す。また、販売結果の分析と考察を行い、提案手法の有効性を検証する。

6.1 実験条件

本実験では ZOZOUSED の 255,350 件のアイテムを対象とし、2020 年 2 月 26 日から 2020 年 3 月 15 日までを出品日として提案手法の検証実験を行った。全出品アイテムをランダムに 2 つのグループに分割し、それぞれ「現行のプロパー期間を適用するアイテム群」（以下、現行）と「提案手法によるプロパー期間を適用するアイテム群」（以下、提案）とした。ただし、出品したアイテム数がそれぞれ均等になるように分割した。現行と提案の各ブランドクラスにおけるアイテム割合を表 8 に示す。また、現行と提案の各アイテムに適用したプロパー期間を表 9 に示す^{*5}。

本実験では、ブランドクラスごとに設定した目標値 α と実測値 $\hat{\alpha}$ の差（目標値 α からの乖離度）を指標とした評価、収益性における評価、管理コストにおける評価の 3 項目について評価を行う。

目標値 α と実測値 $\hat{\alpha}$ の差は、式 (4) に示す目標値 α から

^{*5} 企業の規定により、プロパー期間を操作したため表 7 と若干異なる。

表 8 現行と提案の出品アイテム割合 [%]

Table 8 The rates of items in old/new methods [%].

ブランドクラス	現行のアイテム	提案のアイテム	総アイテム
1	3.01	2.95	2.98
2	0.92	0.91	0.91
3	2.47	2.45	2.46
4	6.79	6.66	6.72
5	23.89	24.21	24.05
6	49.47	49.78	49.62
7	3.51	3.62	3.56
8	9.94	9.42	9.68
計	100.00	100.00	100.00

表 9 現行のプロパー期間と提案のプロパー期間 [日]

Table 9 The old and new proper periods [days].

ブランドクラス	現行	提案 (2月)	提案 (3月)
1	60	39	39
2	60	63	71
3	30	42	37
4	15	29	23
5	15	32	19
6	15	32	18
7	15	15	15
8	15	26	25

の乖離度を用いる。この評価によって、出品者が決定した α % 点がプロパー期間内に販売されたアイテムの割合 (プロパー消化率) に近づいているか、つまり、出品者が意図した販売結果に近づいているかを評価することができる。また、式 (4) におけるプロパー消化率の実測値 $\hat{\alpha}$ は、式 (5) を用いて算出する。

$$\text{目標値 } \alpha \text{ からの乖離度} = |\hat{\alpha} - \alpha| \quad (4)$$

$$\hat{\alpha} = \frac{\text{プロパー期間内に売れたアイテム数}}{\text{販売されたアイテム数}} \times 100 \quad (5)$$

この評価指標を用いて、提案手法によって目標値 α と実測値 $\hat{\alpha}$ の乖離度がどの程度改善されたかを表 10 に示す。ただし、表 10 に示す値は、目標値 α からの乖離度について、「現行のプロパー期間を適用したアイテムで計算した値」から「提案のプロパー期間を適用したアイテムで計算した値」を減算したものである。すなわち、この値が大きければ、現行のプロパー期間に比べて提案のプロパー期間の方が出品者が意図した販売結果に近づいていると解釈することができる。

収益性における評価指標には、式 (6) に示す粗利率を用いる。この評価によって、利益増加の見込みがあるかどうかを評価することができる。ただし、表 11 および表 12 に示す値は、粗利率について「提案のプロパー期間を適用したアイテムで計算した値」から「現行のプロパー期間を適用したアイテムで計算した値」を減算した差を示してい

表 10 各ブランドクラスにおける目標値 α からの乖離度の改善度 (現行 - 提案) [%]

Table 10 Degree of improvement in the difference between parameter α and rate on sold items on proper period for each brand classes [%].

ブランドクラス							
1	2	3	4	5	6	7	8
-0.14	-1.91	3.86	6.57	4.97	3.96	0.03	7.38

表 11 ブランドクラスごとの粗利率の改善度 (提案 - 現行) [%]

Table 11 Degree of improvement in gross profit margin for each brand class [%].

ブランドクラス							
1	2	3	4	5	6	7	8
-1.26	-0.11	1.35	0.85	0.59	0.40	0.21	0.74

表 12 出品月ごとの粗利率の改善度 (提案 - 現行) [%]

Table 12 Degree of improvement in gross profit margin for each listed month [%].

2月	3月
1.19	0.40

る。つまり、この値が大きければ、現行のプロパー期間に比べて提案のプロパー期間の方が利益増加の見込みがあると解釈できる。

$$\text{粗利率} = \frac{\text{合計販売価格} - \text{合計買取価格}}{\text{合計販売価格}} \times 100 \quad (6)$$

管理コストにおける評価指標には、式 (7) に示すアイテム消化率、式 (5) に示すプロパー消化率 $\hat{\alpha}$ を用いる。この評価によって、管理コストが現行に比べて増加しているか否かを評価することができる。ただし、表 13 および表 14 に示す値は、アイテム消化率およびプロパー消化率について、「提案を適用したアイテムにおける値」から「現行を適用したアイテムにおける値」を減算したものとす。つまり、これらの値が大きいくほど、現行のプロパー期間に比べて提案のプロパー期間の方が出品数に対してより多くのアイテムを販売できていると解釈できる。いい換えれば、より管理コストがかかっていないということになる。

$$\text{アイテム消化率} = \frac{\text{販売されたアイテム数}}{\text{出品されたアイテム数}} \times 100 \quad (7)$$

6.2 実験結果と考察

得られた販売結果に対して分析を行い、ブランドクラスごとに設定した目標値 α と実測値 $\hat{\alpha}$ の差 (目標値 α からの乖離度) を指標とした評価、収益性における評価、管理コストにおける評価を行う。分析対象とするアイテムは、前節で述べた出品アイテムのうち、2020年2月26日から2020年10月9日までに販売された237,296件を用いる。

表 13 ブランドクラスごとのアイテム消化率・プロパー消化率の改善度 (提案 - 現行) [%]

Table 13 Degree of improvement in the rate of sold items on all period and proper period for each brand class [%].

	ブランドクラス							
	1	2	3	4	5	6	7	8
アイテム消化率の改善度	-1.62	-0.08	1.08	-0.66	-0.52	-0.11	-0.57	-0.15
プロパー消化率の改善度	-0.14	-1.91	3.86	6.57	4.97	3.96	0.03	4.31

表 14 出品月ごとのアイテム消化率・プロパー消化率の改善度 (提案 - 現行) [%]

Table 14 Degree of improvement in the rate of sold items on all period and proper period for each listed month [%].

	2月	3月
アイテム消化率の改善度	-0.45	-0.24
プロパー消化率の改善度	8.42	3.02

6.2.1 ブランドクラスごとの目標値 α と実測値 $\hat{\alpha}$ に基づく提案手法の評価

5章で述べたプロパー期間の決定において、プロパー期間内で販売したいアイテム割合として事前に設定した目標値 α と実測値 $\hat{\alpha}$ の差 (目標値 α からの乖離度) を指標とした評価を行う。各ブランドクラスにおける目標値 α からの乖離度の改善度 (現行 - 提案) を表 10 に示す。

表 10 より、ブランドクラス 1, 2 は現行の方が目標値 α からの乖離度 (目標値 α と実測値 $\hat{\alpha}$ との差) が小さく、ブランドクラス 3, 4, 5, 6, 7, 8 は提案の方が目標値 α からの乖離度が小さいことが分かる。これらの値は「現行を適用したアイテムにおける値」から「提案を適用したアイテムにおける値」を減算し算出したものであるため、ブランドクラス 4, 5, 6, 7, 8 においては目標値 α と提案のプロパー消化率 $\hat{\alpha}$ との乖離度の方が小さな値であることが分かる。また、9割以上のアイテムがブランドクラス 4, 5, 6, 7, 8 に所属することが事前分析より明らかとなっている。目標値 α はプロパー期間内で販売したいアイテム割合として出品者の意思決定により設定される値であることから、アイテム全体としては、プロパー期間を変更したことにより、出品者の意図に沿った販売結果を得ることができたといえる。

6.2.2 収益性における評価

収益性において販売結果の分析・評価を行う。まず、ブランドクラスごとの収益性について分析する。表 11 に、粗利率についてブランドクラスごとの現行における値と提案における値の差 (提案 - 現行) を示す。

表 11 より、ほとんどのブランドクラスで提案の利益率

が現行の利益率を上回っていることが分かる。約 9 割のアイテムがブランドクラス 4~8 に所属することから、アイテム全体では、提案のプロパー期間を適用することで利益率が增加すると考えられる。

次に、出品月ごとの収益性について分析する。表 12 に、粗利率について出品月ごとの現行における値と提案における値の差 (提案 - 現行) を示す。表 12 より、2月, 3月ともに提案の利益率が現行の利益率を上回っていることが分かる。したがって、2月と3月においては提案のプロパー期間を適用する方が現行に比べて利益率は増加すると考えられる。

以上のことから、提案手法を実運用に適用することで、利益増加の見込みがあると考えられる。

6.2.3 管理コストにおける評価

最後に管理コストにおける販売結果の分析・評価を行う。まず、ブランドクラスごとのアイテム消化率とプロパー消化率の改善度 (提案 - 現行) を表 13 に示す。表 13 より、ほとんどのクラスで提案のアイテム消化率が現行のアイテム消化率を下回っていることが分かる。しかしながら、それらはブランドクラス 1 を除いて 1%未満の値であるため、管理コストにおける大きな差はないと考えられる。一方、プロパー消化率に注目すると、ほとんどのクラスで提案のプロパー消化率が現行のプロパー消化率を上回っていることが分かる。さらに、ブランドクラス 3, 4, 5, 6, 8 は 3%以上の差があることが分かる。約 9 割のアイテムがブランドクラス 4~8 に所属することから、アイテム全体では、提案のプロパー期間を適用することでプロパー消化率が増加したと考えられる。

次に、出品月ごとのアイテム消化率とプロパー消化率の改善度 (提案 - 現行) を表 14 に示す。表 14 より、2月, 3月ともに、提案のアイテム消化率が現行のアイテム消化率を下回っており、プロパー消化率の観点から、提案手法は現行の方法を上回っていることが分かる。また、アイテム消化率の差は 0.5%未満であるため、管理コストにおける大きな差はないと考えられる。一方、プロパー消化率の改善度は 3%以上であるため、2月と3月においては提案のプロパー期間を適用することで、現行に比べてプロパー消化率が増加することが考えられる。

以上のことから、提案のプロパー期間を用いてアイテムを出品することで、全体としての管理コストは従来と大きな変化はないが、値下げをせずに販売されるアイテムの割合が大きくなるといえる。したがって、提案手法の適用により、管理コスト大きく増加させずに出品価格のままアイテムを販売することができると考えられる。

7. 考察

ここでは、アイテムのグルーピング方法、目標値 α の設定、実証実験のデータ分割の観点から、本研究の提案手法について考察を行う。

7.1 アイテムのグルーピング方法について

本研究における提案手法では、アイテムのグルーピングの際にブランドクラスという特徴量を用いた。最適なアイテムのグルーピングという観点においては、より多くの特徴量を考慮し、類似アイテムのグルーピングが可能な、*k*-means [21] や PLSA [22] といったクラスタリング手法も有効であると考えられる。しかしながら、本研究では、決定したプロパー期間の実運用を最終目標としているため、運用コストも考慮する必要がある。上述した手法を適用した実運用を考えると、アイテムが入荷されるたびにクラスタリングを実行しなおし、クラスを決定する必要があると考えられる。一方、本研究で用いたブランドクラスは、あらかじめ分類基準が決定されており、新たなアイテムに対してはその基準に従ってクラス分けを行うだけでよく、簡便な方法となっている。また、アイテムの特徴を最も表している特徴量の1つである価格を考慮して分類を行うため、運用コストも小さく、適切なアイテムのグルーピングが行われているといえる。

このような、価格帯に基づいた分類基準はファッションアイテムに限らず、様々なジャンルのアイテムで用いられることが多く、ビジネスでのシステム運用において有用であると考えられる。本研究では、ZOZOUSUED を対象事例としたが、この分類基準は一般的な2次流通、特に複数ブランドのアイテムを扱うなど、アイテムに多様性がある場合においても有用であることが期待できる。

7.2 目標値 α について

本研究の提案手法では、あらかじめ設定しておく値として目標値 α を設けている。この値は出品者がプロパー期間内に販売したいアイテムの割合として設定するため、現場の状況に応じて検討すべきである。

本研究の実装においては、過去の販売状況に基づいて表6のように決定したが、過去の販売状況は時間経過によって変化することも考えられる。そのため、一定期間経過し、販売状況が変化した際には、プロパー期間を再調整することが必要である。本提案手法は、アイテムの販売状況に応じたモデルの再構築が必要となるものの、 α を変更することで、何度もプロパー期間を決定することができるため、出品者の意図によって多様な決定が可能な手法であるといえる。したがって本提案手法は出品者の意図した販売結果に導くための補助的な役割を果たす手法と考える。

7.3 実証実験のデータ分割について

本研究で行った実証実験では、新たに出品するアイテムをランダムに2等分されるような分割を行った。本来であれば、価格やカテゴリといったすべての特徴量において、数量や構成比が2等分されるような分割を行うべきと考えられる。しかしながら、本研究の対象事例である中古ファッ

ションアイテムは1点ものという性質を持つため、そのような考慮を行うことは容易ではない。そこで、本研究では、アイテムのグルーピングの方法において用いたブランドクラスの構成比が2等分になるような分割を行った。表8から、分割後のアイテムでは、ブランドクラスの構成比がアイテム全体における構成比と大差ないことが分かる。したがって、実証実験におけるアイテムの分割は適切であったと考えられる。

8. まとめと今後の課題

本研究では、ZOZOUSUED におけるアイテムのプロパー期間の決定問題を対象事例とし、新規出品アイテムの出品月に応じた適切なプロパー期間の決定を目的とした、価格帯の類似性を考慮したブランドクラスごとに掲載期間の予測分布を推定するモデルを構築し、出品月を変化させながら推定を繰り返すことで適切なプロパー期間を決定する手法を提案した。また、実際にZOZOUSUED の販売履歴データに適用し、分析することで提案手法の有用性を示すとともに、プロパー期間を決定した。加えて、新規出品アイテムに決定したプロパー期間を適用した検証実験を設計・実施し、実運用における提案手法の有用性を示した。これにより、当該サイトにおいて、より合理的なプロパー期間の決定を可能とするモデルを構築することができた。提案手法により、多様なアイテムを扱うECサイトにおける、より合理的な値下げの仕組みの構築が期待できる。

今後の課題として、提案手法における予測分布の推定モデルの検討があげられる。本研究では、NGBoost を用いて掲載期間の予測分布を推定した。しかしながら、確率分布を推定する手法としては他にもベイズ推定 [23] などが考えられる。これらの分布推定方法を検討することでより適切なプロパー期間の決定が可能となることが期待できる。

謝辞 本研究の一部は JSPS 科研費 21H04600 の助成を受けたものです。

参考文献

- [1] 三好哲也：アパレルオンラインショッピングにおける消費者特性の分析：データ分析コンペティションデータの分析を通して、経営システム, Vol.27, No.2, pp.61–69 (2017).
- [2] 里村卓也：トピックモデルによる顧客データの統合的分析、オペレーションズ・リサーチ, Vol.63, No.2, pp.67–74 (2018).
- [3] 奥野拓也, 中村和幸：状態空間モデルによる購買間隔の規則性の推定、オペレーションズ・リサーチ, Vol.63, No.2, pp.83–90 (2018).
- [4] Kanazawa, S., Sugiyama, Y., Yang, T. and Goto, M.: A Study of Feature Clustering Analysis based on the Hidden Layer Representation of an Autoencoder, *Total Quality Science*, Vol.5, No.1, pp.11–22 (2019).
- [5] Liu, N., Ren, S., Choi, T.M., Hui, C.L. and Ng, S.F.: Sales Forecasting for Fashion Retailing Service Industry: A Review, *Mathematical Problems in Engineering*, Vol.2013, No.738675 (2013).

- [6] Duan, T., Avati, A., Ding, D.Y., Basu, S., Ng, A.Y. and Schuler, A.: NGBoost: Natural Gradient Boosting for Probabilistic Prediction (2019), available from (<https://arxiv.org/abs/1910.03225>).
- [7] 山田紗弥, 朝日弓未: ファッション EC サイトにおける季節を考慮した消費行動分析, 東海大学紀要情報通信学部, Vol.10, No.2, pp.14–19 (2017).
- [8] 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).
- [9] 杉山啓太, 豊田秀樹, 長尾圭一郎, 磯部友莉恵, 岡 律子: ファッション EC サイトにおけるイノベーター検出モデル—基準変数のある多種混合の項目反応モデリング, オペレーションズ・リサーチ, Vol.63, No.2, pp.75–82 (2018).
- [10] Thomassey, S. and Fiordaliso, A.: A hybrid sales forecasting system based on clustering and decision trees, *Decision Support Systems*, Vol.42, No.1, pp.408–421 (2006).
- [11] Ni, Y. and Fan, F.: A two-stage dynamic sales forecasting model for the fashion retail, *Expert Systems with Applications*, Vol.38, No.3, pp.1529–1536 (2011).
- [12] Au, K.-F., Choi, T.-M. and Yu, Y.: Fashion retail forecasting by evolutionary neural networks, *International Journal of Production Economics*, Vol.114, No.2, pp.615–630 (2008).
- [13] Choi, T.-M., Yu, Y. and Au, K.-F.: A hybrid SARIMA wavelet transform method for sales forecasting, *Decision Support Systems*, Vol.51, No.1, pp.130–140 (2011).
- [14] 仁ノ平将人, 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販売価格予測モデルに関する一考察, 情報処理学会論文誌, Vol.60, No.4, pp.1151–1161 (2019).
- [15] 益子 宗, シルメンバータル・ミヤグマルスレン, 酒巻隆治: KiTeMiROOM: モバイル端末のためのファッションコーディネート支援システム, 電子情報通信学会論文誌, Vol.J96-D, No.10, pp.2286–2294 (2013).
- [16] 田村 悠, 吉住宗朔, 福永 峻, 三宅聡一郎, 片山翔太, 中田和秀: ファッション EC サイトにおけるアンケートを用いたブランド推薦システム, オペレーションズ・リサーチ, Vol.63, No.2, pp.91–98 (2018).
- [17] Friedman, J.H.: Greedy function approximation: A gradient boosting machine, *The Annals of Statistics*, Vol.29, No.5, pp.1189–1232 (2001).
- [18] Amari, S.: Natural Gradient Works Efficiently in Learning, *Neural Computation*, pp.251–276 (1998).
- [19] Park, H., Amari, S. and Fukumizu, K.: Adaptive Natural Gradient Learning Algorithms for Various Stochastic Models, *Neural Networks*, Vol.13, No.7, pp.755–764 (2000).
- [20] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: LightGBM: A Highly Efficient Gradient Boosting Decision Tree, *Advances in Neural Information Processing Systems* (2017).
- [21] MacQueen, J.: Some methods for classification and analysis of multivariate observations, *Proc. 5th Berkeley Symp. Math. Statist. and Prob.*, Vol.1, pp.281–297 (1967).
- [22] Hofmann, T.: Probabilistic Latent Semantic Analysis, *Proc. UAI '99*, pp.289–296 (1999).
- [23] Hastie, T.: *The Elements of Statistical Learning*, Springer (2008).



桑田 和

1996年生。2019年早稲田大学創造理工学部経営システム工学科卒業。2021年同大学大学院創造理工学研究科経営システム工学専攻修了。2021年日本電気株式会社入社。データサイエンティストとして機械学習を用いたデータ分析によるビジネス課題の解決に従事。



三川 健太 (正会員)

1981年生。2005年武蔵工業大学環境情報学部環境情報学科卒業。2007年同大学大学院修士課程修了。2016年早稲田大学大学院博士後期課程修了。博士(工学)。2013年早稲田大学助手。2016年湘南工科大学工学部情報工学科講師。2019年同大准教授。機械学習とその応用に関する研究に従事。IEEE, 電子情報通信学会, 日本経営工学会等各会員。



佐々木 北都

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博士課程修了. 博士(工学). 1997年同大学理工学部助手. 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.