

A Decision Making Model for Second-hand Fashion Items
Based on Sales History Data

KUWATA Izumi

1. 研究背景と目的

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 における値下げのタイミングの適正化に関する問題を対象とする。ZOZOUSED は、ユーザーから古着を買取り、中古品として再販売する事業を展開している。その際、出品する商品（以下、アイテム）の売れ残りを防ぐため、出品から一定期間経過したアイテムを段階的に値下げする仕組みを採用している。アイテムを出品した際の初期価格は出品価格と呼ばれ、収益性の観点からは値下げをせずに出品価格で販売されることが望ましいが、在庫コストや出品アイテムの鮮度といった観点からは、値下げを許容し早く販売することも必要となる。そのため、出品価格を保持する期間の適切な設定は重要な課題の一つとなっており、本研究ではこの保持期間を「プロパー期間」と定義する。

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ーの意思決定によるものであり、各アイテムの出品から誰かに購入されるまでの期間（以下、掲載期間）はアイテムの購入されやすさを表したものと考えられる。そのため、プロパー期間の決定には予想される掲載期間を考慮する必要がある。しかし、「掲載期間の正確な一点予測は困難である」、「季節ごとに掲載期間の傾向が異なる」という2つの課題がある。これらの解決のため、前者に対しては掲載期間の予測分布を推定することを考え、後者に対してはアイテムの出品月を考慮した予測モデルを構築することを考える。

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、各アイテムの掲載期間を確率分布で予測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)[1] を用いる。さらに、アイテムの特徴量のうち、出品月のみを変化させて予測を繰り返すことで、各月の適切なプロパー期間を決定する。本研究では、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ことでその有効性を検証する。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、販売結果に対する実証評価と考察を行う。

2. 対象問題

2.1. 問題設定

上述のように、ZOZOUSED では、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ているが、収益性の観

点からは出品価格のまま購入されることが望ましい。しかし、管理コストの観点においては、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、ZOZOUSED における従来のプロパー期間は経験的に決定されているため、これまで値下げされて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。そのため、過去の販売履歴データを有効に活用し、客観的にも適切なアイテムのプロパー期間を設定することが望まれている。

ZOZOUSED では幅広いアイテムを扱っているため、アイテムの多様性を考慮した適切なプロパー期間の決定が必要である。加えて、ユーザーの意思決定が反映されたアイテムの掲載期間を考慮する必要がある。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、以下の2つの問題点が存在する。

- 各アイテムの掲載期間は様々な要因による不確実性が非常に高く、単純な回帰モデルによる一点予測は困難である。
- 季節性のあるアイテムも多く、出品する時期によって掲載期間の傾向が異なる可能性がある。

これらを考慮した掲載期間の予測モデル構築とプロパー期間の決定方法が必要であり、本研究ではその具体的な方法を提案する。まず、前者に対しては、掲載期間の予測確率分布を推定することを考える。後者に対しては、出品月が掲載期間に寄与する特徴量とみなし、出品月ごとにプロパー期間を設定することを考える。

2.2. 事前分析

本節では、2018年1月1日から2018年12月31日に販売されたアイテムの掲載期間に関する基礎分析を行い、その特徴を把握する。図1に掲載期間におけるアイテム数を示す。

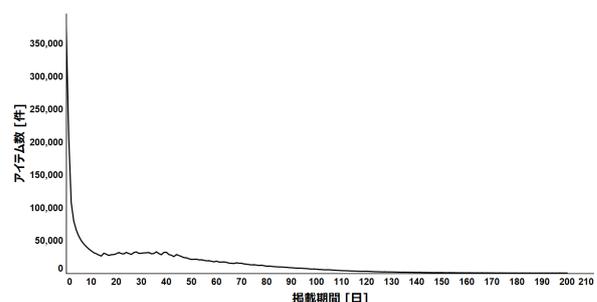


図 1: 各掲載期間におけるアイテム数

図 1 に示すように、掲載期間は 0 日が最も多くなっており、急速に減少してから 10~50 日程度で平坦となり、また減少するような形状となっている。つまり、本サイトで扱われるアイテム多くは出品初日に販売されており、その一方で、販売まで 1 年近くかかるアイテムも存在していることが分かる。このことから、出品初日以外で販売されたアイテムの掲載期間はばらつきが大きいことが言える。

3. Natural Gradient Boosting

本研究では、掲載期間の予測確率分布を推定可能な手法として、Natural Gradient Boosting (NGBoost) の適用を考える。

NGBoost は勾配ブースティングをベースとした手法であり、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 つの構成要素を持つ。これらに対し、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 [2] 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。

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

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。

4. 提案手法

4.1. 着想

前述の通り、ZOZUSED 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、合理的ではない。そこで、類似したアイテムごとに特徴量との関係性を把握するため、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管理単位である「ブランドクラス」を用いてアイテムをグルーピングする。ブランドクラスとは、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。販売単価はアイテムの類似性に大きく寄与する特徴量であると考えられる。そのため、ブランドクラスを用いることで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。

前述のように、各アイテムの掲載期間はばらつきが大きく、掲載期間そのものを正確に一点予測することは困難である。また、事前分析より、掲載期間の分布は全体的な概形では指数分布に近いことが明らかとなっている。そのため提案手法では、3 章で述べた指数分布を用いた NGBoost により各アイテムの掲載期間の予測確率分布を推定する。これにより、

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。

また、掲載期間には季節性が存在し、出品月も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みなし、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用いて適切なプロパー期間を決定することを考える。

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、学習データに対して NGBoost を適用し、各アイテムを出品した際の掲載期間の分布予測モデルを構築する。このとき、予測確率分布には指数分布を仮定する。

次に、各アイテムの特徴量のうち、出品月を 1 月から 12 月まで変化させながら、テストデータに対して構築した予測モデルを適用することを繰り返す。テストデータの出品月を変化させる度に予測モデルを適用し、各アイテムをその月に出品した場合の掲載期間の予測確率分布を得る。得られた分布から、 α % 点の掲載期間を算出する。ただし、 α は値下げをせずに販売したいアイテムの割合として、実応用の観点から現場レベルで決定される値である。そして、ブランドクラスごとに、各アイテムの予測確率分布の α % 点にあたる掲載期間の平均値を適切なプロパー期間とする。

4.3. 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。 N 個のアイテムのうち、 n 番目のアイテムの特徴量を $\mathbf{x}_n \in \mathbb{R}^D$ 、掲載期間を $y_n \in \mathbb{R}$ とする。ここで、 D は特徴量数とする。ブランドクラスは前述の通り 8 クラスとし、 M 個のアイテムのうち、各ブランドクラスに属するアイテム数を $M_i (i = 1, \dots, 8)$ 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通りである。

Step1) 掲載期間の予測確率分布を推定

各アイテムの特徴量 \mathbf{x}_n を説明変数、掲載期間 y_n を目的変数として、NGBoost により、ブランドクラスごとの予測確率分布を推定する。

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) M_i 個のアイテムと出品月 j に対して、得られた分布の α % 点の掲載期間 $s_m^{ij} (1 \leq m \leq M_i)$ を算出する。

Step2-5) M_i 個のアイテムの s_m^{ij} の平均値を算出し、プロパー期間 t_{ij} とする。

Step2-6) $j = 12$ のとき, Step2-7 へ. $j < 12$ のとき, $j = j + 1$ として Step2-3 へ.

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.

5. 実データ分析

本章では, ZOZOUSED において過去の販売履歴データに提案手法を適用して, 提案手法の有用性について検証するとともに, 適切なプロパー期間を決定する.

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. 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 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. 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¹. 掲載期間の予測モデルの構築においては, 出品価格, 定価などの計 28 要素からなる特徴量 ($D = 28$) を用いた.

5.1. 分析 1: 予測モデルの有用性の検証

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, 他の予測モデルとの比較を行う.

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一つである, LightGBM[3] を用いる. この手法は区間予測ができないため, 評価指標は比較手法に合わせて平均平方二乗誤差 (RMSE) を用いる.

予測モデルによる予測掲載期間の RMSE を表 1 に示す.

表 1: 予測モデルに対する 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 1 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.

5.2. 分析 2: プロパー期間の決定

次に, 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. NGBoost の弱学習器の数は実験的に適切な値を調べ, 500 と設定した. また, NGBoost で用いる予測のための確率分布には指数分布を仮定した.

¹非公開情報のため詳細は記載できないが, 買取方法の変化ならびにデータ取得方法の関係でテストデータが多くなっている.

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 は, 現行のプロパー期間内での販売実績を基に「プロパー期間内に販売したいアイテム割合」としてブランドクラスごとに設定した. 設定した α の値を表 2 に示す.

表 2: ブランドクラスごとに設定した α [%]

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

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 3 に示す. ただし, ここでは 1 月から 6 月までのプロパー期間を示す.

表 3: 決定したプロパー期間 [日]

| 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 |

表 3 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. 前述の通り, ブランドクラスは主に平均販売単価が高い順にグルーピングしている. ZOZOUSED では, 価格帯が高いアイテムほど, 長くサイトに掲載され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であると考えられる.

以上のことから, 実運用可能なプロパー期間決定モデルが構築できたと言える.

6. 提案手法を用いた実運用における実証実験

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実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実証実験を行った. 全出品アイテムを, アイテム数が均等になるようにランダムに 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「新しいプロパー期間を適用するアイテム群」(以下, 提案) とした. 現行と提案の各アイテムに適用したプロパー期間を表 4 に示す. ただし, ZOZOUSED での規定²により, 前章で決定したプロパー期間と若干異なる. なお, 現行は出品月によらず一定である.

本実験では, ブランドクラスごとに設定した α 点に関する評価と収益性に関する評価を行う.

ブランドクラスごとに設定した α 点に関する評価指標には, 式 (2) に示す α との乖離度を用いる. この評価によ

²プロパー期間の下限日数が存在し, それを下回った期間は適用できない.

表 4: 現行のプロパー期間と提案のプロパー期間 [日]

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

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に近づいているか、すなわち、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 (2) におけるプロパー消化率は、販売されたアイテム数に対するプロパー期間内に売れたアイテム数の割合として算出する。

$$\alpha \text{ からの乖離度} = |\alpha - \text{プロパー消化率}| \quad (2)$$

収益性における評価指標には、式 (3) 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。

$$\text{粗利率} = \frac{\text{合計販売価格} - \text{合計買取価格}}{\text{合計販売価格}} \times 100 \quad (3)$$

6.2. 実験結果と考察

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 件を分析対象アイテムとし、実証実験の結果を示す。

6.2.1. ブランドクラスごとに設定した α に関する評価

5 章のプロパー期間の決定において事前に設定した α に対して、販売結果がどの程度近づいているかについて評価を行う。この α はプロパー期間内で販売したいアイテム割合としてサイト側が戦略的に決定する値であるため、評価には式 (2) を用いる。ただし、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、新しい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、その差分によって評価する。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新しい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。ブランドクラスごとに算出した値 (現行 - 提案) を表 5 に示す。

表 5: α からの乖離度の改善度 (現行 - 提案) [%]

| 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 5 より、ブランドクラス 1, 2 は現行の方が α からの乖離度が小さく、ブランドクラス 3~8 は提案の方が α からの乖離度が小さいことが分かる。また、9 割以上のアイテムがブランドクラス 4~8 に所属することが事前分析より明らかとなっている。 α はプロパー期間内で販売したいアイテム割合としてサイト側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、サイト側の意図に沿った販売結果を得ることができたと言える。

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。ここでは、式 (3) を用いてブランドクラスごとの収益性について分析す

る。ただし、新しいプロパー期間を適用したアイテムで計算した粗利率から、現行のプロパー期間を適用したアイテムで計算した粗利率を減算し、その差分を算出して評価する。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新しいプロパー期間の方が利益増加の見込みがあると解釈することができる。ブランドクラスごとに算出した値 (提案 - 現行) を表 6 に示す。

表 6: 粗利率の改善度 (提案 - 現行) [%]

| 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 6 より、ほとんどのブランドクラスで提案の粗利率が現行の粗利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 4~8 に所属することから、アイテム全体では、新しいプロパー期間を適用することで粗利率増加が期待でき、実際にアイテム全体における粗利率は 0.58% 増加した。

一方で、平均販売単価が高いブランドクラス (特にブランドクラス 1) においては、提案の方が粗利率が低くなっていることが分かる。これは、アイテムを現行と提案に分割する際、アイテムの買取単価及び出品単価を均等に分けることが出来なかったことも原因であると考えられる。

以上より、提案手法を実運用することで、全体としては利益増加の見込みがあると言える。

7. まとめと今後の課題

本研究では、ZOZOUSED における、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手法の確立を目的とした。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確率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 ZOZOUSED の販売履歴データに適用し、分析結果を通じて提案手法の有用性を示した。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実証実験を設計・実施することで、実証的に提案手法の有用性が示された。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。

今後の課題として、本研究では、NGBoost を用いて掲載期間の予測確率分布を推定したが、ベイズ推定を用いた方法も考えられる。加えて、提案手法の掲載期間の予測確率分布には、指数分布を仮定したが、ワイブル分布を適用することで掲載期間の予測精度の向上が期待でき、今後の課題とする。

参考文献

- [1] Duan, T., Avati, A., Ding, D. Y., Basu, S., Ng, A. Y., Schuler, A., "NGBoost: Natural Gradient Boosting for Probabilistic Prediction," <https://arxiv.org/abs/1910.03225>, 2019.
- [2] Amari, S., "Natural Gradient Works Efficiently in Learning," *Neural Computation*, Vol. 10, No. 2, pp.251-276, 1998.
- [3] 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," *In Advances in Neural Information Processing Systems*, 2017.