

# 売買履歴の分析による古着売買の価格査定 支援モデルの提案

## A Price Assessment Support Model for Secondhand Clothing Sales by Trading History Analysis

高島悠佑<sup>1\*</sup> 後藤裕介<sup>1</sup> 南野謙一<sup>1</sup> 渡邊慶和<sup>1</sup>

Yusuke Takashima<sup>1</sup>, Yusuke Goto<sup>1</sup>, Ken'ichi Minanimino<sup>1</sup>, and Yoshikazu Watanabe<sup>1</sup>

<sup>1</sup> 岩手県立大学ソフトウェア情報学部

<sup>1</sup> Faculty of Software and Information Science, Iwate Prefectural University

**Abstract:** In secondhand clothing trading, the same item can have different values depending on the status, color and size. Price assessment is carried out with reference to tacit knowledge and auction trading information. We propose a price assessment support model that predicts a selling price by inputting the used clothes' information, and recommends an appropriate selling price and buying price. We analyze historical trading data, and use a machine learning algorithm to predict possible selling prices. We develop an formula of evaluating trading history that assures a proper profit from trading. We analyze historical trading data to assess whether they have problems in profit by using the proposed model.

### 1. はじめに

近年、エコロジー意識から多くの人が着られなくなった衣類をリサイクルするためアパレル系のリサイクル業が注目を集めている。

古着の売買では1点ごとに状態・色・サイズが異なり、同じ商品でも価値が変化するため、スタッフの経験に基づく“暗黙知[1]”やオークションサイトの約定結果などの相場を参考にして行われている。単一商品に注目して売価分析をする関連研究は多数存在するが、安定した供給を前提としており、同じ商品が存在しない古着売買に適用することは難しい。

本研究では、同一商品が安定して供給されることを前提とせず、買取時の商品や環境情報を入力すると将来の販売価格を予測し適切な売出価格と買取価格を推薦する価格査定支援モデルを提案する。買取と販売の実績データ(POS データ)を用いて、データ期間中に買取・販売が行われ、重量買取やセット販売が行われていない取引に着目して分析を行い、機械学習により販売価格を予測する。取引の望ましさの評価式を開発し、予測された販売価格から望ましさを一定水準以上にする買取・売出価格を推薦する。提案モデルは実際の取引データを比較し評価する。

### 2. 分析対象の概要

アパレル系リサイクル業でフランチャイズ事業により全国に約 90 の店舗を持つ企業よりデータ提供を受けて研究を行った。各店舗で買取と販売を行うため、店舗により商品の需要や供給が異なり、価格査定を行う商品にも違いが生じる。買取価格は洋服の人気度によって異なり有名ブランドなどは1点ごとに査定を行い、店舗ごとに特に人気の高いものは1kgあたり500円で買取としており店舗により価格が変動する(表1)。また、特徴として販売価格を10段階の価格帯で設定しており、特定の曜日に価格帯が1段階下がることが挙げられる。そのため、多く

表 1: 買取査定の基準と設定価格

基準	設定価格
有名ブランドの洋服	1点ごとに査定
非常に人気の洋服	
特に人気の洋服	1kgあたり500円
人気の洋服	1kgあたり50円
その他 (店舗ごと異なる)	1kgあたり10円
	1kgあたり1円
	1kgあたり1円

\* 連絡先: 岩手県立大学ソフトウェア情報学部  
〒020-0693 岩手県滝沢市菓子 152-52  
E-mail: g031m091@s.iwate-pu.ac.jp

の商品に値下がり幅が存在する。この値下がり幅の大小は売出価格に大きく左右され、最大幅は最小幅の10倍の値下がり幅となっており、売出価格は利益を予測する上で非常に重要なものとなる。

今回提供いただいたデータは、2015年5月1日～2017年4月30日までの買取実績データ約418万件と販売実績データ約835万件である。実績データには、店舗、日付、担当者、価格、数量、ブランド、ジャンルなどの情報が含まれている。しかし、売出日付、色、サイズ、傷み具合など衣類の詳細に関する情報は含まれていない。また、2017年の5月と8月の2回にわたり行ったインタビュー調査から、店舗ごとにPOSデータの管理をしており、登録されていないブランドやジャンルの追加などは店舗単位で行なわれているため、店舗により登録名や登録コードが異なるものが複数存在していることが分かった。さらに、企業全体を統合して管理するためのマスタなどが存在していないことや買取と販売を結びつけるユニークコードが存在しておらず店舗と在庫コードを組み合わせる必要があることなどが判明した。

### 3. 価格査定に関連する研究

価格査定に関連する研究として、不動産などの価格査定の研究が多数されている。加藤[2]の研究では、多店舗のPOSデータを用いて食料関係の消耗品の価格査定を行う研究がされている。この研究では、競合商品の価格と価格の大台に着目し比較を行っている。商品の価格はメーカーな違いなどに左右されること50円や100円などの大台の有無が購買傾向の変化に関係することが明らかにされている。

また、荒川ら[3]の研究では飲食店向き不動産の賃料に関する価格査定を機械学習で予測するモデルの研究がされている。不動産の持つ様々な情報を基に暗黙知の予測と学習器による精度の比較が行われている。

上記の研究と本研究の対象には需要や供給、価格の大きさなど異なる点が存在し、そのままでは古着売買に適用することはできない。しかし、大台差の有無や機械学習による価格推定などの考え方は本研究の対象とする古着売買でも参考になるため、2つの研究の考え方を取り入れ古着売買に適したモデルを構築する。

## 4. 価格査定支援の提案モデル

### 4.1 提案モデルの概要

本研究は、売買履歴をもとに買取店舗・ブランド・

ジャンルなどの情報を入力することで販売価格・売出価格・買取価格の目安を出力することで利益損失のリスク軽減を目的とした価格査定を支援するモデル提案を行う。図1は提案モデルの流れを整理したものである。取引情報は、企業から提供された2年分の買取実績データと販売実績データをもとにデータ期間内に一点ごとに査定して買取から販売まで行われたもの対象とする。この取引情報をもとに取引の評価式を示して評価を行う。その後、ランダムフォレスト回帰で販売価格の予測を行い、売出価格の価格帯と評価式の値が一定以上になる買取価格の推薦をして出力をする。

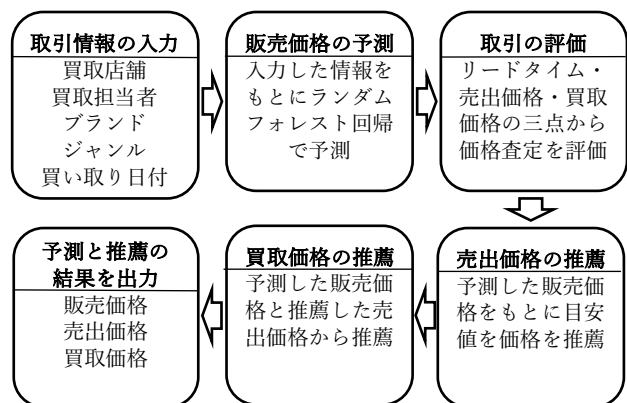


図1：提案モデルの流れ

### 4.2 データの準備

ランダムフォレストで予測を行うにあたり、今回提供されたデータのクリーニング、結合、絞り込みを行った。

はじめに、クリーニングとして、提供されたデータにおいて買取と販売のデータの結合を行うために必要な在庫コードが入力されていないものと取引数量が複数のものの削除を行った。

結合に際して、結合のためのユニークなコードが存在しないことから、店舗と在庫コードを組み合わせたものを結合キーとして結合を行い、686,712件のデータが抽出された。

最後に、分類の大きいジャンルの服飾雑貨や仮入力のコードなど、商品の特定が難しいもの、対象企業で設定されている10段階の価格と売出価格が一致していないもの、値下がり幅に異常がなく買取から販売までの日数が70日以下のものに絞ることで、99,987件の実績データが得られた。

しかし、店舗によるコードと項目名の不一致があり、統一は現状困難であると判断したため一部コードと項目が一致していないものが含まれている。

### 4.3 販売価格の予測

価格査定支援を行うにあたり実際に査定を行う従業員が必要とする情報の中で特に重要であり経験や暗黙知を必要とするものは何かと考えたとき、需要や供給などの大きく作用され企業側が予測することが困難な販売価格を予測することで買取価格や売出価格の目安を導き出しやすくなるのではないかと考え、機械学習を用いて予測することにした。

販売価格を予測するために、機械学習の1つであるランダムフォレスト回帰を用いることにした。販売価格を目的変数として、POSデータの管理が店舗でされていることや地域差を考慮するために「店舗」、商品の需要や人気度などに直結する「ブランド」、「ジャンル」、従業員による査定結果の違いなどの暗黙知の考慮をするために「買取担当者」、季節や時期などの考慮をするために「買取日付」の5項目を説明変数に設定して学習を行った。

この時の学習サンプル数は抽出したデータの約50%に当たる49,978件として学習を行った。パラメータの最適化をするために、はじめに説明変数の数による精度の比較を行った(図2)。決定木の説明変数の数が2のとき、平均2乗誤差が最小で重相関の2乗が最大であったため、説明変数の数を2に設定した。次に説明変数の数が2のときの決定木の数による精度の比較を行った(図3)。決定木の数が500以降は平均2乗誤差、重相関の2乗ともに大きな差はなかったため、決定木の数を500に設定した。

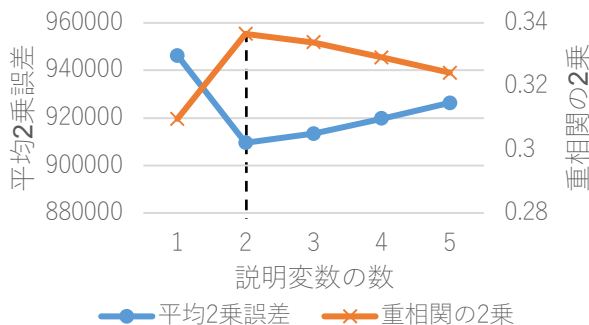


図2：説明変数の数による精度の比較

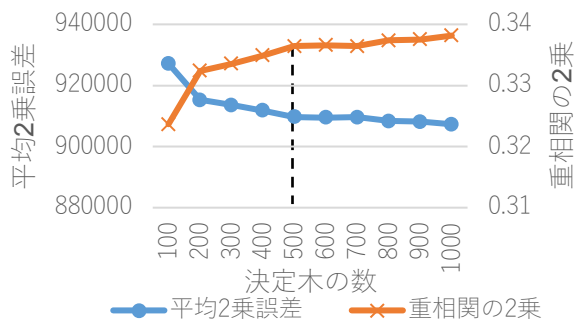


図3：決定木の数による精度の比較

この時の平均2乗誤差は909,657.5で重相関の2乗は0.3364384であった。誤差の範囲が非常に大きいため誤差範囲の縮小が今後の課題と考えられる。

### 4.4 価格査定の評価式

図4は、買取から販売までの流れを図で表現したものである。買取をA、売出をB、販売をCとし、それぞれの価格をp日付tで表現している。本研究では、販売価格 $p'_i$ と買取価格 $p_i$ の差( $p'_i - p_i$ )を利益、売出価格 $p_i^*$ と販売価格 $p'_i$ の差( $p_i^* - p'_i$ )を値下がり幅、買取日から販売日までの期間( $t'_i - t_i$ )をリードタイムとする。

以上をふまえ、取引における価格査定を価格pと日付tを用いて定量的に評価するために、値下がり幅、利益、リードタイムの3点から評価を行う以下の式を開発した。

$$D_i = DS_i + PF_i + LT_i \quad (1)$$

式(1)はある商品iの取引の値下がり幅の評価 $DS_i$ 、利益の評価 $PF_i$ 、リードタイムの評価 $LT_i$ の総合的な評価 $D_i$ である。この値の最大値は3で値が大きいほど望ましい取引ができていていることを表す。

$$DS_i = 1 - \frac{(p_i^* - p'_i)}{p_i^*} \quad (2)$$

式(2)は値下がり幅の評価 $DS_i$ である。販売価格 $p'_i$ が売出価格 $p_i^*$ に近い(値下がり幅が小さい)ほど、高い評価とする。そのため値下がり幅( $p_i^* - p'_i$ )を値下がりの最大値(売出価格 $p_i^*$ )で割り(値下がりの割合)、1から引くことで値下がり幅の評価式とした。

$$PF_i = \frac{p'_i - p_i}{p_i^*} \quad (3)$$

式(3)は利益の評価であり、利益が大きいほど、評価は高く、利益がマイナスの場合は評価もマイナスの値となる。利益が最大の利益(売出価格 $p_i^*$ )に近いほど、予定通りの利益を得られていると考え、高い評価とするため利益を売出価格 $p_i^*$ で割ることで利益の評価式とした。

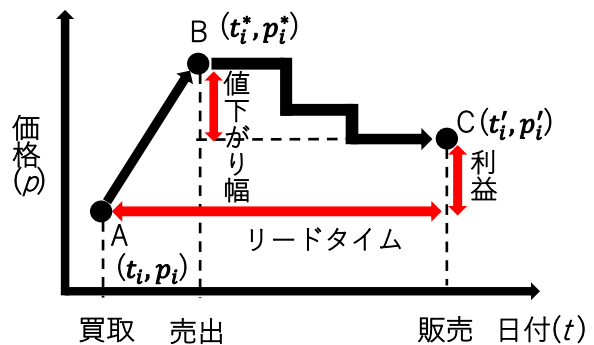


図4：買取から販売までの流れ

$$LT_i = \frac{t_{max} - (t_i' - t_i)}{t_{max}} \quad (4)$$

式(4)はリードタイムの評価 $LT_i$ であり、取引の最大日数 $t_{max}$ とリードタイムの関係を表した式である。今回のデータでは正確な売日付 $t_i'$ がわからないためリードタイムの始点を買日付 $t_i$ とし、販売日付 $t_i'$ までの経過日数をリードタイムとしている。リードタイムが短いほど、高い評価とする。

この評価式をもとに販売価格 $p_i'$ をランダムフォレスト回帰で求め、売価格 $p_i'$ や一定の割合の利益が見込める買価格 $p_i$ を推薦することで、暗黙知を使わず価格査定への支援をすることが可能となると考える。

#### 4.5 売価格と買価格の推薦

売価格 $p_i'$ は割引やセールなどの値下がり考慮し、予測した売価格 $p_i'$ を分析対象の企業により定められている10段階の価格帯を参考にして予測した売価格よりも1段階上の価格帯を推薦する。

買価格 $p_i$ は、4.4の式(3)を $p_i = p_i' - (PF \times p_i')$ と変形し予測した売価格 $p_i'$ と推薦した売価格 $p_i'$ からPFの値が一定基準を満たすような買価格 $p_i$ を推薦する。ただし買価格 $p_i$ の下限は0円とする。そのため、PFの目標値を設定する必要がある。

買価格 $p_i$ を推薦するときの目標値を決めるためにPFの目標値に仮の値を代入したときに実績の価格とどの程度一致しているのか買価格帯ごとにランク分けし比較を行った(表2)。この比較から、幅の大きいランク10を除いた実績と推薦が同じランクに属するものの総数に対する割合、実績と推薦が同じか±1のランクに属するものの総数に対する割合を求めグラフにした(図5)。グラフからPFの値に0.65を代入したときの一一致割合が0.58および0.22と最も高かったため、0.65をPFの目標値することで、実績値に近い買価格 $p_i$ を推薦できると考えられる。

表2：買価格のランク分け

ランク	価格
10	1500円より高い
9	1300円より高く1500円以下
8	1100円より高く1300円以下
7	900円より高く1100円以下
6	700円より高く900円以下
5	500円より高く700円以下
4	300円より高く500円以下
3	100円より高く300円以下
2	0円より高く100円以下
1	0円

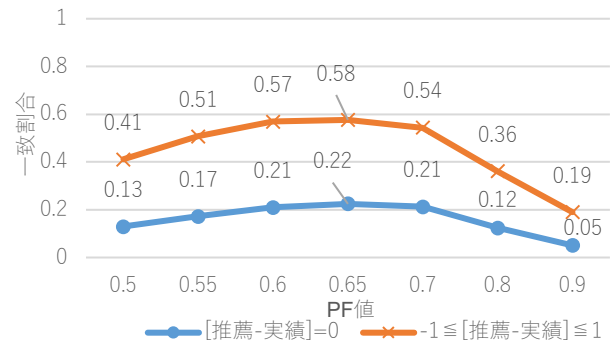


図5：PFの値による推薦と実績の一一致割合

### 5. 提案モデルを使った分析

#### 5.1 推薦と実績の差の分析

売価格を10段階の表をもとにランク付けし、さらに買価格の10段階に分類しランク付けしたものを比較した。この比較結果をもとに、推薦と実績の差が小さいものを(i)望ましい取引、推薦よりも実績の方が大きく予定通りの利益を得られなかったものを(ii)望ましくない取引、買価格が推薦よりも実績の方が大きく利益が推薦結果よりも小さくなってしまったものを(iii)買取で失敗した取引、推薦よりも実績の方が小さいものを(iv)安く見積もりした取引、売価格が推薦よりも実績の方が大きいものを(v)売出で失敗した取引と5つに分類した(表3)。この時の望ましい取引は予測させた取引の約5割を占めている一方で望ましくない取引が約2割あり、取引の半分は望ましい取引が実現されているものの、残りの半分は商品価値の予測に失敗し十分な利益を得られなかった取引が存在していることが分かった。

表3：取引の分類ごとの件数と割合

分類	件数	割合
望ましい	25,767	0.52
望ましくない	10,731	0.22
買取で失敗	5,875	0.12
安く見積もり	4,743	0.10
売出で失敗	2,894	0.06

#### 5.2 望ましい取引の商品に関する分析

望ましい取引の件数が特に多い店舗に注目し望ましい取引に多く含まれるブランドとジャンルの組み合わせを調べた(表4)。その結果、件数の多い上位10の組み合わせに同じブランドや同じジャンルの商品

が多く、望ましい取引に含まれるブランドやジャンルには偏りがあるのではないかと考え、特に件数の多かったブランド「RAGE BLUE」とジャンル「長袖はおり」に注目して調べた(表 5)。

望ましい取引に多く含まれるブランド「RAGE BLUE」は、望ましい取引が約 6 割ある一方、買取で失敗したものや実績よりも推薦の方が高く商品の価値を全体的に安く見積もってしまったものの割合もたかくなっている。望ましい取引に多く含まれるジャンル「長袖はおり」は、望ましい取引が 5 割ある一方、望ましくない取引も 3 割近くある。そのため、取引においてブランドとジャンルの組み合わせが商品価値に対する影響が大きく、価格査定においてブランドとジャンルの組み合わせを考慮し価格査定を行うことが重要であることが分かった。

表 4：望ましい取引の多い組み合わせ上位 10 件

ブランド	ジャンル	望ましい	総数	割合
RAGE BLUE	長袖はおり	32	44	0.73
NOBRAND	BAG	28	32	0.88
NIKO and	パンツ	23	28	0.82
BROWNY	長袖はおり	22	26	0.85
NIKO and...	ワンピース	22	28	0.79
earth music&ecology	スカート	19	20	0.95
NIKO and...	長袖はおり	18	22	0.82
USED	長袖はおり	18	34	0.53
Crisp	ワンピース	17	17	1.00
Ank Rouge	ワンピース	16	9	0.64

表 5：望ましい取引が多いブランドとジャンル

RAGE BLUE の取引について			長袖はおりの取引について		
分類	件数	割合	分類	件数	割合
望ましい	83	0.58	望ましい	378	0.50
望ましくない	9	0.06	望ましくない	203	0.27
買取で失敗	18	0.13	買取で失敗	74	0.10
安く見積もり	22	0.15	安く見積もり	62	0.08
売出で失敗	10	0.07	売出で失敗	35	0.05
合計	142	1.00	合計	752	1.00

## 6. まとめ

本研究では、古着売買を行う企業の価格査定の手助けを行うためのモデルの提案を行った。本研究のモデルは、買取から販売まで行われた商品の実績データを学習させランダムフォレスト回帰により販売価格の予測を行うことで、売価や買取価格の目安

となる数値の推薦を行うものである。本研究のモデルの特徴として、実績データを扱うことで暗黙知を必要とせずに販売価格を予測し売価・買取価格の目安を求めることが可能なことである。モデルを現場で使用するには誤差範囲が大きすぎるため安定した結果が得られず困難と考えられるが、学習器の精度向上による誤差範囲の縮小をすることで実用可能になると考える。

そのため、今後の課題として学習器による精度の比較、予測精度の向上、誤差範囲の縮小が挙げられる。また今回使用したデータには、商品の色やサイズなど取引した衣類の詳細な情報が含まれていないため、それらの情報を追加することで精度の向上が期待できる。

## 謝辞

研究を行うにあたりデータ提供等のご協力いただいた企業の皆様に深く感謝いたします。

## 参考文献

- [1] 野中郁次郎, “知識創造企業とデザイン,” デザイン学研究特集号, Vol.7, No.2, pp.2-17 (2000)
- [2] 加藤弘貴, “多店舗 POS データを利用した売価分析モデルの適用方法の検討,” 流通情報, Vol.478, pp.23-31 (2009)
- [3] 荒川周造, 諏訪博彦, 小川祐樹, 荒川豊, 安本慶一, 太田敏澄, “暗黙知に基づく飲食店向け不動産賃料推定モデルの提案,” 情報処理学会論文誌, Vol.59, No.1, pp.33-42 (2018)