

## 論文・事例研究

# ファジィクラスタワイズ回帰を用いた 共同購入型クーポンサイトの閲覧傾向分析

高野 祐一, 田中 未来, 鮎川 矩義, 竹山 光将, 神里 栄, 千代 竜佑, 小林 健,  
田中 研太郎, 中田 和秀

## 1. はじめに

共同購入型クーポンサイトとは、商品やサービスの大幅割引や特典のついたクーポンを、期間限定で、一定数以上の申し込みがあった場合にのみ販売するウェブサイトである。ユーザはクーポンを購入することで大幅割引や特典を利用でき、クライアントはユーザにサービスを利用してもらうことで店舗の稼働率を上げることができる(図1)。さらに、クーポンの販売には一定数以上の申し込みが必要となるため、申し込みをしたユーザがクーポンを購入するためにインターネット上で自ら販売促進を行うという利点がある。また、必要な申し込み数を損益分岐点を上回るところに設定すれば、クーポンの販売によって必ず三者に利益が生まれるといった利点もある。

今回分析対象とした共同購入型クーポンサイトでは、トップページに多くのクーポンが表示されている。しかしながら、これらのクーポンはすべてのユーザに対して一律に表示されているために、ユーザによっては全く興味のないクーポンが表示されてしまう可能性がある。商品販売によって売上が発生するようなサイトでは、ユーザが迷わずに商品購入ページに遷移できるようにサイト内のコンテンツやレイアウトなどを改善する導線最適化が重要であり、ユーザ別に対策を行うことが有効である[8]。ゆえに共同購入型クーポンサイトにおいても、ユーザの導線に適切なクーポンを配置することでサイトの売上を改善できる可能性があり、そのためには各ユーザの興味を把握することが必要となる。

たかの ゆういち, たなか みらい, すけがわ のりよし,  
たけやま みつまさ, かみざと さかえ, ちよりゆうすけ,  
こばやし けん, たなか けんたろう, なかた かず  
ひで

東京工業大学 大学院社会理工学研究科 経営工学専攻  
〒152-8552 東京都目黒区大岡山 2-12-1

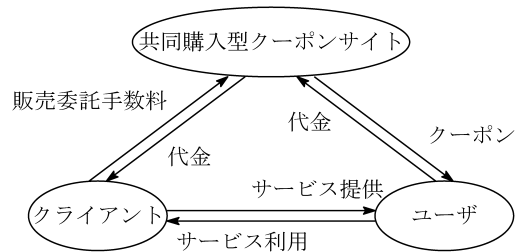


図1 共同購入型クーポンサイトのビジネスモデル

各ユーザの興味を把握するためにユーザ一人ひとりに対して予測モデルを作成することが考えられる。しかし、推定に十分なサンプル数を確保できない場合が多く、ユーザ管理の手間なども考慮すると適切な方法とは言えない。このため、ユーザの興味・嗜好を反映してユーザを適切に分類し、分類されたユーザ群に対して予測モデルを構築するようなアプローチが有効だと考える。

本論文では、クーポンの各ジャンルに対するPV(ページ閲覧)数をユーザの興味を表す指標と定め、ファジィクラスタワイズ回帰[1, 4, 13]を利用してPV数予測モデルを作成する。クラスタワイズ回帰とは、異なるサンプルが混合されたデータセットに対して、予測精度が向上するようにサンプルのクラスタリングと各クラスタの回帰モデルの作成を同時に行う手法である[2, 6, 11]。この際に、各クラスタへの帰属度(メンバーシップ値)に曖昧さ(複数のクラスタに所属すること)を許すモデルがファジィクラスタワイズ回帰であり、理論・応用の両面から多くの研究がなされている[3, 5, 7, 10, 12]。また、ファジィクラスタワイズ回帰には分類の曖昧度を調節するパラメータが用意されており、通常のクラスタワイズ回帰の拡張となっている。

本研究では、まず曖昧度パラメータとクラスタ数を変化させながら予測精度を調べ、分類に曖昧さを許す

ことで予測精度が向上することと、予測精度の観点からするとクラスタ数は4もしくは7で十分であることを確認した。次に、ほかの手法と予測精度を比較することでファジィクラスタワイズ回帰の有効性を検証した。さらに、抽出された4つのクラスタの特徴を多面的に分析することで、実際のマーケティング施策につながるような有益な知見を得ることができた。

本論文は以下のように構成される。2節では、分析データの概要を説明する。3節では、ファジィクラスタワイズ回帰の定式化と解法を説明する。4節では提案モデルの予測精度を検証し、5節では分析結果に基づく施策の提案を行う。最後に6節でまとめと今後の課題を述べる。

## 2. 分析データ

本研究では経営科学系研究部会連合協議会主催、平成24年度データ解析コンペティションで提供されたデータを利用した。提供された2011年7月1日から2012年6月30日までの1年間のデータを4カ月ごとに3つに分割し、以下のような2つのデータセットを作成した。提供されたデータにおける全ユーザ数は22,873名であり、各データセットにおける分析対象は1年間を通して会員だったユーザのうち、訓練用データの期間における総PV数が上位30%のユーザとした。なお、1年間を通して会員だったユーザに対して、それぞれのデータセットで分析対象としたユーザの1年間の購入金額の割合はそれぞれ54.6%と60.3%であった。また、2つのデータセットで重複している分析対象ユーザ数は2,445名であった。

データセット1 (分析対象ユーザ数: 3,290名)

訓練用データ

2011年7月1日から2011年10月31日まで

検証用データ

2011年11月1日から2012年2月28日まで

データセット2 (分析対象ユーザ数: 3,328名)

訓練用データ

2011年11月1日から2012年2月28日まで

検証用データ

2012年3月1日から2012年6月30日まで

今回提供されたデータのうち、クーポンの総数は19,723件であり、「グルメ」、「宅配」、「健康・医療」、「その他」など13種類のクーポンジャンルに分類されている。クーポンジャンル「宅配」に分類されるクーポンの総数は5,930件と、全クーポンのうち約30%を占めていた。そこで本研究ではクーポンジャンル「宅配」

に分類されているクーポンを、9種類のクーポンジャンルに再分類した。詳しい分類については、図5を参照されたい。なお、クーポンジャンル「健康・医療」に分類されているクーポンの総数は59件と少なかったため、これらのクーポンは取り除いて分析を行った。

## 3. ファジィクラスタワイズ回帰

本節では、Wedel and Steenkamp [13]に従ってファジィクラスタワイズ回帰の定式化と解法を説明する。

### 3.1 回帰式

特定のジャンルに対するPV数は、そのジャンルに属するクーポンの数に強く依存する。したがって、ユーザの各ジャンルへの興味を把握することを目的とした場合、単純なPV数よりも、1クーポンあたりの平均PV数に着目するほうが適切だと考えられる。そこで、本研究では、ユーザ $u \in \mathcal{U}$ のクーポンジャンル $g \in \mathcal{G}$ に対する1クーポンあたりの平均PV数を被説明変数 $Y_{u,g}$ とする。すなわち、

$$Y_{u,g} = \frac{\text{ユーザ } u \text{ のジャンル } g \text{ に対する PV 数}}{\text{ジャンル } g \text{ のクーポン数}}$$

と定義する。このために、PV数 $Y_{u,g}$ は必ずしも整数値とはならないことに注意されたい。

クーポンジャンルからPV数を予測するために、クーポンジャンル $g \in \mathcal{G}$ に対して説明変数 $i \in \mathcal{I} \subset \mathcal{G}$ の値を

$$X_{i,g} = \begin{cases} 1 & (i = g \text{ のとき}) \\ 0 & (i \neq g \text{ のとき}) \end{cases}$$

と定義する。これらの説明変数を用いたクラスタ $c \in \mathcal{C}$ の回帰式は以下ようになる：

$$Y_{u,g} = a_{c,0} + \sum_{i \in \mathcal{I}} a_{c,i} X_{i,g} + e_{c,u,g} \quad (\forall c \in \mathcal{C}, \forall u \in \mathcal{U}, \forall g \in \mathcal{G}). \quad (1)$$

ただし、 $a_{c,0}$ は切片、 $a_{c,i}$ は偏回帰係数、 $e_{c,u,g}$ は残差とし、これらの値は各クラスタ $c \in \mathcal{C}$ で異なる。例えばクーポンジャンル $g=2$ に着目すると、定義より $X_{2,2} = 1$ 、 $X_{1,2} = X_{3,2} = X_{4,2} = \dots = 0$ となる。このとき、クーポンジャンル2に対する回帰式(1)は $Y_{u,2} = a_{c,0} + a_{c,2} + e_{c,u,2}$ となる。これはクーポンジャンル2に対するPV数 $Y_{u,2}$ を切片 $a_{c,0}$ とクーポンジャンル2の偏回帰係数 $a_{c,2}$ で予測することを意味する。

■積乗型の回帰式 被説明変数であるPV数 $Y_{u,g}$ は

常に非負の値をとるが、回帰式 (1) では予測値が非負となる保証はない。また、説明変数の効果は足し算の形ではなく掛け算の形で現れると想定すると、以下の積乗型の回帰式（例えば、文献 [9] を参照）が有効となる：

$$Y_{u,g} = \exp(a_{c,0} + e_{c,u,g}) \prod_{i \in \mathcal{I}} X_{i,g}^{\alpha_{c,i}} \quad (\forall c \in \mathcal{C}, \forall u \in \mathcal{U}, \forall g \in \mathcal{G}). \quad (2)$$

回帰式 (2) は両辺の対数を取り、 $Y_{u,g} \leftarrow \log Y_{u,g}$ 、 $X_{i,g} \leftarrow \log X_{i,g}$  とすることで回帰式 (1) に帰着できる。この方法で推定した回帰式の有効性は 4 節で検証する。

### 3.2 定式化

ファジィクラスタワイズ回帰の定式化を示す。まず、ユーザ  $u$  のクラスタ  $c$  への帰属度を表すメンバシップ値を  $z_{c,u}$  とし、メンバシップ値は以下の条件を満たすとする：

$$0 \leq z_{c,u} \leq 1, \quad (\forall c \in \mathcal{C}, \forall u \in \mathcal{U}), \quad (3)$$

$$\sum_{c \in \mathcal{C}} z_{c,u} = 1, \quad (\forall u \in \mathcal{U}). \quad (4)$$

クラスタワイズ回帰では、制約条件 (1), (3), (4) の下で、メンバシップ値  $z_{c,u}$  で重み付けられた残差 2 乗和が最小になるように、各ユーザのメンバシップ値と各クラスタの回帰式のパラメータを決定する [6]：

$$\left| \begin{array}{l} \text{最小化} \\ \alpha_{c,0}, \alpha_{c,i} \\ e_{c,u,g}, z_{c,u} \\ \text{制約条件} \end{array} \right. \sum_{c \in \mathcal{C}} \sum_{u \in \mathcal{U}} z_{c,u} \sum_{g \in \mathcal{G}} e_{c,u,g}^2 \quad (5)$$

問題 (5) は、通常のクラスタワイズ回帰に対して、同じユーザのサンプル  $(u, g)$  は同じクラスタに分類するという制約が課されたモデルとみなせる。この最適化問題を解くと、特別な場合を除いてメンバシップ値  $z_{c,u}$  は 0 か 1 に貼りつくことが知られている [6]。このとき、ユーザ  $u$  は回帰式の残差  $\sum_{g \in \mathcal{G}} e_{c,u,g}^2$  の値が最も小さくなるようなクラスタ  $c$  に割り振られる。

一方で、ファジィクラスタワイズ回帰では、以下のように分類の曖昧度を表すパラメータ  $\mu (> 1)$  を導入する [13]：

$$\left| \begin{array}{l} \text{最小化} \\ \alpha_{c,0}, \alpha_{c,i} \\ e_{c,u,g}, z_{c,u} \\ \text{制約条件} \end{array} \right. \sum_{c \in \mathcal{C}} \sum_{u \in \mathcal{U}} z_{c,u}^\mu \sum_{g \in \mathcal{G}} e_{c,u,g}^2 \quad (6)$$

曖昧度パラメータの値を  $\mu \downarrow 1$  とした極限では問題 (6)

は問題 (5) と等価になる。一方で曖昧度パラメータ  $\mu$  の値を大きくしていくと、徐々に一人のユーザが複数のクラスタに所属するようになり、 $\mu \uparrow \infty$  とするとすべてのユーザ  $u$  について各クラスタに対するメンバシップ値を等しく  $z_{c,u} = 1/|\mathcal{C}|$  とすることが問題 (6) の最適解となる。

### 3.3 解法

問題 (6) は目的関数が非凸非線形の最適化問題であり、厳密解を求めることは困難である。そのため、加重最小 2 乗法によるクラスタごとの回帰分析と、Lagrange の未定乗数法によるメンバシップ値の最適化を交互に繰り返す発見的な解法が提案されており [1, 4, 13]、本研究でも同様の解法を利用した。

## 4. 予測精度の検証

本節では、2 節で説明したデータにファジィクラスタワイズ回帰を適用し、その予測精度を検証する。

なお、予測精度を検証する際には、まず訓練用データを用いて問題 (6) を解いた。次に、検証用データにおいて、各クラスタの回帰式による予測値をメンバシップ値で重み付けして予測値を計算し、検証用データの決定係数を求めた。本節では、2 つのデータセットにおける検証用データの決定係数の平均値を示して予測精度を評価する。

### 4.1 積乗型回帰の有効性

まず、今回のデータに対する積乗型回帰 (2) の有効性を簡単に確認する。全ユーザを一括にして予測した場合、素朴な線形回帰の決定係数が 0.04 と非常にあてはまりが悪かったのに対し、積乗型回帰の決定係数は 0.14 と 3.5 倍の差があった。また、ファジィクラスタワイズ回帰の場合も同様の結果が得られた。

### 4.2 パラメータの決定

次に、曖昧度パラメータ  $\mu$  とクラスタ数  $k = |\mathcal{C}|$  の変動が、どのように回帰の結果に影響するのかを考察する。

図 2 は、各  $\mu \in \{1.001, 1.1, 1.2, 1.3, 1.5, 2.0\}$  と各  $k \in \{1, 2, \dots, 10\}$  についてファジィクラスタワイズ回帰を行い、得られた回帰式の決定係数を計算したものである。

曖昧度パラメータ  $\mu$  については、ほとんど曖昧さを許さない場合 ( $\mu = 1.001$ ) よりも、少しだけ曖昧さを許す場合 ( $\mu = 1.1, 1.2, 1.3$ ) のほうがよい結果が得られた。よって、今回対象としたデータに対しては、通常のクラスタワイズ回帰のモデルよりもファジィクラスタワイズ回帰のモデルが適していると言える。しか

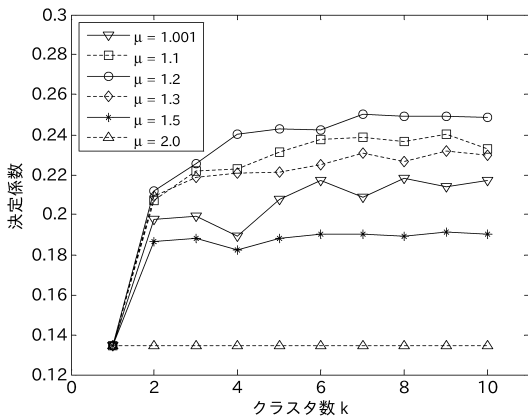


図2 曖昧度パラメータ  $\mu$  とクラスタ数  $k$  の変化に伴う決定係数の変化

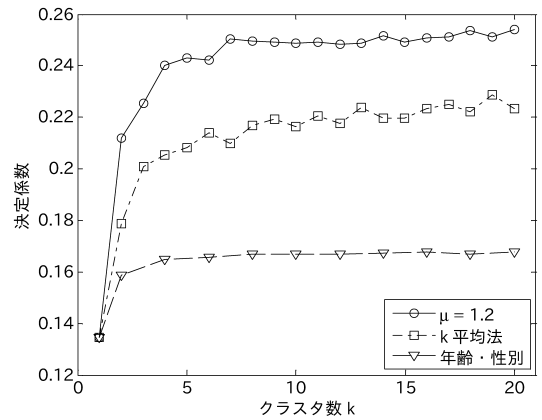


図4 他の手法との比較

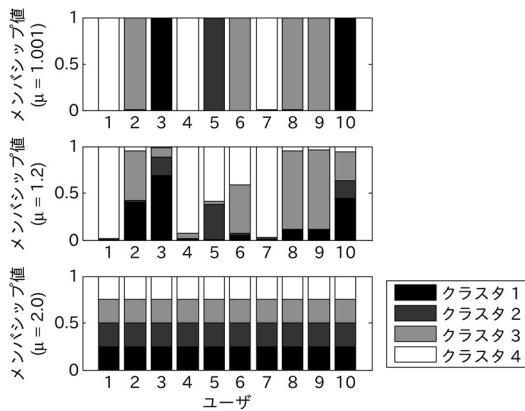


図3 曖昧度パラメータ  $\mu$  の変化に伴うメンバシップ値の変化

しながら、 $\mu$  の値を大きくすればするほど精度が向上するというわけではなく、 $\mu = 1.2$  で頭打ちになっている。

次に、クラスタ数について考察する。クラスタ数  $k$  を増やせば訓練用データに対してはあてはまりのよいモデルが得られるが、過剰適合が生じて検証用データにおいては予測精度が悪化する可能性がある。図2から、 $\mu = 1.2$  の場合は  $k$  としては4あるいは7で十分であると言える。

図3は、 $\mu = 1.001, 1.2, 2.0$  の場合のデータセット1における一部のユーザのメンバシップ値を図示したものである。 $\mu = 1.001, 1.2$  の場合には、メンバシップ値がユーザによって大きく異なるのに対し、 $\mu = 2.0$  では、そのような個人差が失われていた。以降では、 $\mu = 1.2$  の場合についてより深く考察していく。

#### 4.3 他の手法との比較

クラスタワイズ回帰の利点は、予測精度が向上する

ように分類を行う点であったが、実際にその効果があったかを検証する。ここでは、性別と年齢による分類と、各クーポンジャンルに対するPV数<sup>1</sup>を用いた  $k$  平均法による分類をそれぞれ行った後で各クラスタに対して積乗型回帰を行ったものを比較対象とする。図4からわかるように、ファジィクラスタワイズ回帰を用いることで、素朴な分類手法では得られないような予測に適した分類を得ることができた。

### 5. クラスタの解釈と施策の提案

最後に、データセット1において  $\mu = 1.2, k = 4$  としたときに得られた4つのクラスタを考察し、具体的なマーケティング施策の例を述べる。

各クラスタの偏回帰係数の値を図5に示す。ただし、0.1%有意とならなかったクラスタ2の「グルメ」とクラスタ3の「宅配/日用品」の偏回帰係数は図5から除いてある。なお、多重共線性を回避するためにクーポンジャンル「その他」は説明変数から削除した。このために、偏回帰係数の値はすべて「その他」との相対値となっており、偏回帰係数が正の場合は「その他」と比較してPV数が多く、負の場合は「その他」と比較してPV数が少ない傾向があると解釈できる。

まず、「ギフトカード」はどのクラスタでも偏回帰係数の値が高く、多くのユーザが高い反応度を示すクーポンであることがわかる。クラスタ1は「レジャー」、

<sup>1</sup>  $k$  平均法では、各クーポンジャンルのPV数を平均0、標準偏差1に正規化して使用した。

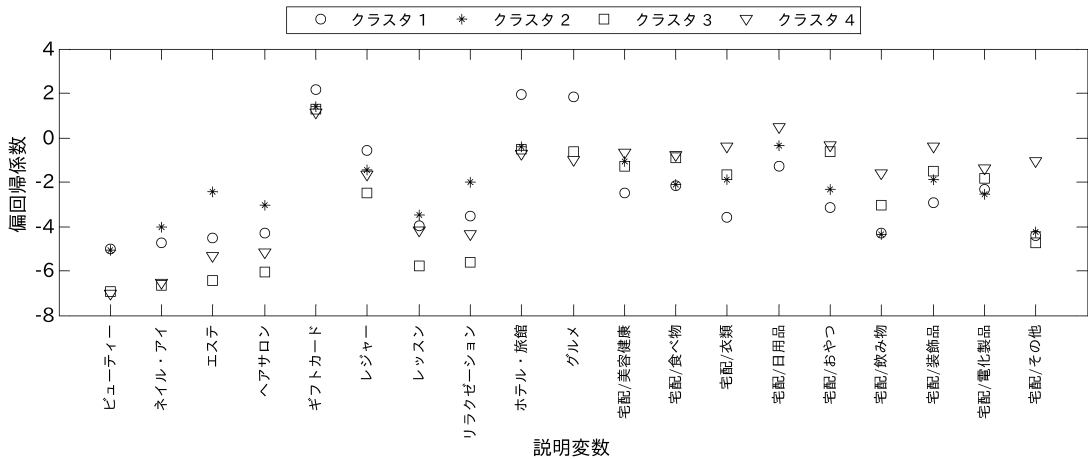


図5 各クラスタの0.1%有意となった偏回帰係数

タとする。クラスタ2は「ビューティー」、「ネイル・アイ」、「エステ」、「ヘアサロン」、「リラクゼーション」など美容に関するクーポンへの反応度が高く、このクラスタを「美容好き」クラスタとする。クラスタ3は「ネイル・アイ」、「エステ」、「ヘアサロン」、「レジャー」、「レッスン」、「リラクゼーション」など外出を伴うようなクーポンに対する反応度が低く、このクラスタを「インドア派」クラスタとする。クラスタ4は宅配のクーポン全般に対する反応度が高く、このクラスタを「宅配好き」クラスタとする。なお、 $k=7$ の場合にも、 $k=4$ の場合と同様の特徴を示すクラスタが得られた。

表1は各クラスタの決定係数とメンバシップ値の総和および性別と居住地方の分布を、図6は各クラスタの年齢の分布を示したものである。クラスタ2はほかと比較して決定係数の値が0.21と低く、PV数の予測が難しいクラスタであると言える。各クラスタのメンバシップ値の総和は各クラスタのユーザ数に対応する値であり、それぞれのクラスタに属するユーザの人数に大きなばらつきはないことがわかる。また、表2は各クラスタのユーザー一人あたりの閲覧と購買に関する数値を示したものである。ここで、CV数は商品購買回数を表し、CV率はCV数/アクセス回数とした。また、表2の数値は2011年7月1日から2012年6月30日までの1年間で集計した。以降では、表1、2と図6に基づいて各クラスタの特徴をさらに解析し、各クラスタのユーザに対して有効な施策の例を示す。

■クラスタ1: 「外出好き」クラスタ このクラスタに属するユーザは男性が多く、全体の分布から比べると、60代や70代といった年齢の高いユーザが多かつ

表1 各クラスタの決定係数とメンバシップ値の総和および性別・居住地方の分布

クラスタ	1	2	3	4
決定係数	0.39	0.21	0.47	0.44
メンバシップ値の総和	866.5	656.7	978.5	788.3
男性	66.0%	21.8%	49.7%	30.6%
女性	34.0%	78.2%	50.3%	69.4%
北海道・東北	3.8%	4.6%	6.1%	5.4%
北関東	2.2%	1.0%	3.5%	2.0%
南関東	37.0%	46.3%	34.3%	36.6%
北陸・東山	3.1%	1.7%	5.2%	2.9%
東海	4.6%	4.9%	6.6%	8.6%
近畿	23.7%	21.3%	20.3%	24.6%
中国・四国	2.8%	1.0%	7.8%	5.7%
九州・沖縄	9.6%	4.1%	9.9%	9.1%
登録なし	13.1%	15.1%	6.3%	5.2%

表2 各クラスタのユーザー一人あたりの閲覧と購買

クラスタ	1	2	3	4
アクセス回数	43.9	80.3	50.3	61.6
滞在時間 [秒]	308.0	439.8	312.7	352.0
PV数	97.0	247.3	119.6	164.6
CV数	3.9	7.2	4.8	6.0
CV率	8.9%	9.0%	9.5%	9.7%
購入金額 [千円]	23.8	35.2	23.7	24.7

た。また、アクセス回数、PV数、CV数はほかのクラスタに比べて少ないという傾向が見られた。よって、このクラスタに対しては、中高年の男性を想定したホテルやグルメのクーポンを推薦し、PV数、CV数を高めていくことが必要だと考えられる。



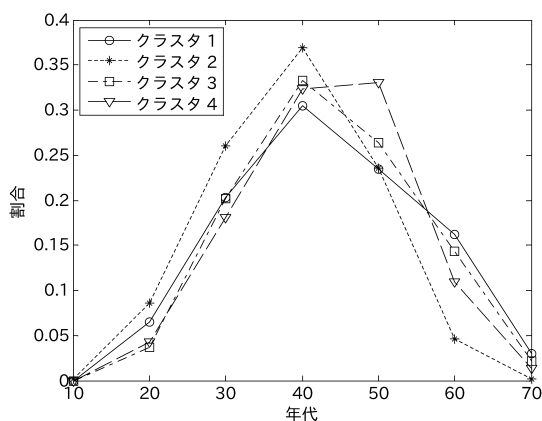


図 6 各クラスタの年齢の分布

■**クラスタ 2: 「美容好き」クラスタ** このクラスタに分類されるユーザの約 8 割は女性で、20 代から 40 代までのユーザが多い。居住地方で見ると、「南関東」と「登録なし」のユーザが多いという傾向が見られた。アクセス回数、PV 数、滞在時間、購入金額がほかのクラスタに比べて高かったため、優良ユーザの集まったクラスタと言える。また、これらの情報から、このクラスタのユーザは吟味して購入を検討する傾向があることがわかる。そのため、サイト内の導線や広告を工夫して見やすいサイトを設計することや、各クーポンの商品情報を充実させることが優良ユーザを増やし、さらにはサイトの売上増加にもつながると考えられる。

■**クラスタ 3: 「インドア派」クラスタ** このクラスタには、男性と女性が半々の割合で存在しており、年齢の分布に特徴的なところはなかった。しかし、居住地方を見てみると、「北海道・東北」、「北陸・東山」、「中国・四国」、「九州・沖縄」などといった地方在住のユーザが多いという特徴がある。また、CV 率は高いが、CV 数も PV 数も少ないため、必要なときに買うものがある程度絞ってサイトに訪問して購入するユーザが多いと考えられる。この結果は、このサイトには地方在住のユーザに向けた外出を伴うクーポンが不足していることを示唆している可能性があり、地方向けのクーポンを充実させることが PV 数、CV 数を増加させることにつながると考えられる。

■**クラスタ 4: 「宅配好き」クラスタ** このクラスタに分類されるユーザは女性が多く、ほかのクラスタに比べて 50 代が多いという特徴が見られた。CV 率など閲覧と購買に関する数値は比較的高く、準優良ユーザと言える。よって、中年女性を想定した宅配関連のクーポンを推薦することが有効だと考えられる。

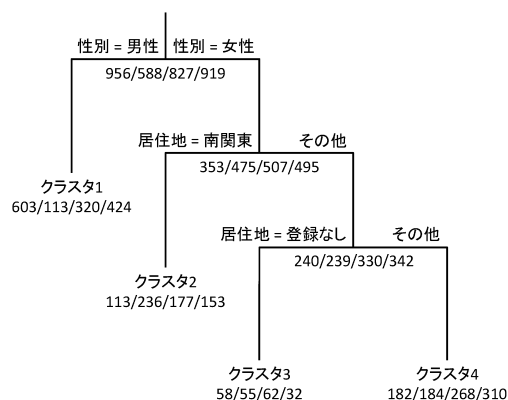


図 7 デモグラフィック情報を用いた決定木

■**EC サイトのレイアウトへの応用** ファジィクラスタワイズ回帰は EC サイトのレイアウトをユーザごとにカスタマイズすることにも利用できる。例えば、得られたメンバシップ値から各ユーザが興味を持つであろうクーポンを推測し、そのようなクーポンの詳細ページへのリンクをトップページの目立つ場所に配置すれば、サイトの PV 数、CV 数の増加につながると考えられる。

■**新規ユーザへの対応** 今回の分析は、当該サイトへの PV 数が一定数以上あるユーザを対象としている。そのため、新規ユーザに対しては PV 数が蓄積されるまでは、会員登録時に得られるデモグラフィック情報（性別、年齢、居住地など）を用いて、そのユーザが属するクラスタを予測することが考えられる。性別、年齢、居住地の情報から、各ユーザをメンバシップ値最大のクラスタに分類する決定木<sup>2</sup>を図 7 に示す。なお、図 7 はデータセット 1 の結果であり、各クラスタのユーザ数を「クラスタ 1/クラスタ 2/クラスタ 3/クラスタ 4」の形式で記載している。この決定木の正答率は 36.8% であり、高くない。これは、デモグラフィック情報だけでは得られないような分類がファジィクラスタワイズ回帰によって得られたことを示している。

## 6. おわりに

本論文では、多様なユーザの興味を考慮しつつクーポンジャンルに対する PV 数を予測するためにファジィクラスタワイズ回帰を用いて分析を行った。今回の解析では、4 つの解釈しやすいクラスタを抽出することができ、また、それらに対するメンバシップ値に応じ

<sup>2</sup> データ解析ソフト R の rpart パッケージを利用し、分割の規準として Gini 係数を使用した。

てユーザへの個別対応を可能にする施策を提案することができた。

今回の解析に用いたファジィクラスタワイズ回帰は汎用性が高く、類似した問題に対しても同様の解析を行うことができる。実際、ファジィクラスタワイズ回帰は、複数の回帰モデルを仮定し、予測精度が高まるようなクラスタリングと各クラスタに対応する回帰モデルの偏回帰係数を求めるという方法であり、複数種類のサンプルが混合されたデータセットについて予測モデルを構築する際に有効である。

ファジィクラスタワイズ回帰分析と似た分析方法として、潜在クラス回帰分析が挙げられる。ファジィクラスタワイズ回帰分析では、曖昧度パラメータを用いてメンバシップ値の乱雑さを制御する点が、潜在クラス回帰分析との大きな相違点である。ファジィクラスタワイズ回帰分析と潜在クラス回帰分析のいくつかの推定手法との比較については今後の課題としたい。

#### 参考文献

- [1] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson, "Detection and Characterization of Cluster Substructure I. Linear Structure: Fuzzy  $c$ -Lines," *SIAM Journal on Applied Mathematics*, **40**(2), 339–357, 1981.
- [2] R. A. Carbonneau, G. Caporossi and P. Hansen, "Globally Optimal Clusterwise Regression by Mixed Logical-Quadratic Programming," *European Journal of Operational Research*, **212**(1), 213–222, 2011.
- [3] P. D'Urso, R. Massari and A. Santoro, "A Class of Fuzzy Clusterwise Regression Models," *Information Sciences*, **180**(24), 4737–4762, 2010.
- [4] R. J. Hathaway and J. C. Bezdek, "Switching Regression Models and Fuzzy Clustering," *IEEE Transactions on Fuzzy Systems*, **1**(3), 195–204, 1993.
- [5] C. C. Kung and J. Y. Su, "Affine Takagi-Sugeno Fuzzy Modelling Algorithm by Fuzzy  $c$ -Regression Models Clustering with a Novel Cluster Validity Criterion," *IET Control Theory & Applications*, **1**(5), 1255–1265, 2007.
- [6] K. Lau, P. Leung and K. Tse, "A Mathematical Programming Approach to Clusterwise Regression Model and Its Extensions," *European Journal of Operational Research*, **116**(3), 640–652, 1999.
- [7] C. Li, J. Zhou, X. Xiang, Q. Li and X. An, "T-S Fuzzy Model Identification Based on a Novel Fuzzy  $c$ -Regression Model Clustering Algorithm," *Engineering Applications of Artificial Intelligence*, **22**(4), 646–653, 2009.
- [8] 小川卓, 「入門 ウェブ分析論」, ソフトバンククリエイティブ, 2010.
- [9] 岡太彬訓, 守口剛, 木島正明 (編), 「マーケティングの数理モデル」, 朝倉書店, 2001.
- [10] M. Soltani, A. Chaari and F. Ben Hmida, "A Novel Fuzzy  $c$ -Regression Model Algorithm Using a New Error Measure and Particle Swarm Optimization," *International Journal of Applied Mathematics and Computer Science*, **22**(3), 617–628, 2012.
- [11] H. Späth, "Algorithm 39: Clusterwise Linear Regression," *Computing*, **22**(4), 367–373, 1979.
- [12] T. Tan, H. W. Suk, H. Hwang and J. Lim, "Functional Fuzzy Clusterwise Regression Analysis," *Advances in Data Analysis and Classification*, **7**(1), 57–82, 2013.
- [13] M. Wedel and J.-B. E. M. Steenkamp, "A Fuzzy Clusterwise Regression Approach to Benefit Segmentation," *International Journal of Research in Marketing*, **6**(4), 241–258, 1989.