

個人の好みに関する感性のモデル化に関する研究

A Study on *KANSEI* Modeling of Preference about Items

荻野 晃大

Akihiro Ogino

京都産業大学 コンピュータ理工学部, 京都市 北区 上賀茂本山

Kyoto Sangyo University, Motoyama Kamigamo, Kita-Ku, Kyoto

あらまし: 本論文では、個人適応型のショッピングサイトを実現するために、商品に関するユーザの好みをモデル化する方法を提案する。本論文で提案するモデルは、ユーザが商品に対する好みを決定する過程を、(1) 商品の各属性を認知する、(2) 商品の各属性について好みの態度を決める、(3) 各属性に関する好みの態度を連合して、最終的な好みの態度を決定すると捉え、それに基づいてユーザの注目している商品の属性・属性値の抽出、抽出した属性・属性値に対する好みの態度の数量化を行い、ユーザが商品の好みを判断する基準となるモデルを作成した。

Summary: This paper proposes a model of user profile for a personalized e-commerce site. The model of user profile estimates attributes and their values of products that are related to user preference. The model of user profile also calculates the degree of relation with the attribute values of product and the user preference. The result shows that we can detect the attribute value related to user preference.

キーワード: 感性モデル, 好み, 個人化, Eコマース

Keywords: *KANSEI* Modeling, Preference, Personalization, E-Commerce

1. はじめに

現在のショッピングサイトには、ユーザが一見ただけでは全ての商品を見られない程の商品が登録されている。そのため、商品名などの商品に関する明確な情報を持ち得ないユーザは、自分の好みに適する商品をショッピングサイトから手早く見つられなくなっている。そのような状況の中で、ユーザの好みに基づいて商品を選別したり、推薦したりすることで、ユーザの商品検索を支援する個人適応型の情報サービス(以後、個人化情報サービス)が注目を浴びている。個人化情報サービスは、ショッピングサイトにおけるユーザの検索・購買履歴やユーザへのアンケートなどの情報から、商品に関するユーザの好みのモデル化を行う。そして、各ユーザのモデルを用いてユーザの好みに適する商品の推定を行う。

好みのモデル化手法には、大きく分けてベクトルを用いる方法、ネットワークを用いる方法、ルールを用いる方法がある。ベクトルを用いる方法は、商品の特徴をベクトルで表現し、ユーザの好きな商品と嫌いな商品の特徴を表すベクトルの違いを SVM(Support Vector

Machine)等により機械学習することで、ユーザの好みに適する商品を推定するモデルを生成する。ネットワークの方法は、ユーザの好き・嫌いと商品の特徴との間の関係をページアンネットワークなどによりネットワークとして形成し、確率的にユーザの好みに適する商品を推定するモデルを生成する。ルールを用いる方法は、アソシエーションルールやラフ集合など用いてユーザの購買履歴等をデータマイニングし、そこから抽出したルールを用いて推薦する方法である。

ベクトルとネットワークを用いる方法は、ユーザの好みを推定する方法として有効であるが、モデルの構造が複雑になる。そのためユーザは、モデルによって推定された結果に対して善し悪しを判断する事はできるが、「モデル自体がユーザの好みをどのように解釈しているのか?」ということを直感的に理解することは難しい。一方、ルールを用いる方法は、IF-THEN形式としてモデルを表現する。そのためユーザは、モデルの内容を直感的に理解できる。しかし、複数のルールが生成された場合には、ユーザが全てのルールの内

容を網羅して、モデルがどのように自分を理解しているのかを理解することは難しい。

本論文では、ショッピングサイトを利用する各ユーザが、個人化情報サービスにより推定された商品の理由と、その元となっている好みのモデルの内容を直感的に理解できる好みのモデル化手法を提案する。

2. 好みのモデル化

本研究では、ユーザがある商品に関する好みの態度を決定する過程を、図1に示すような過程として捉えている。

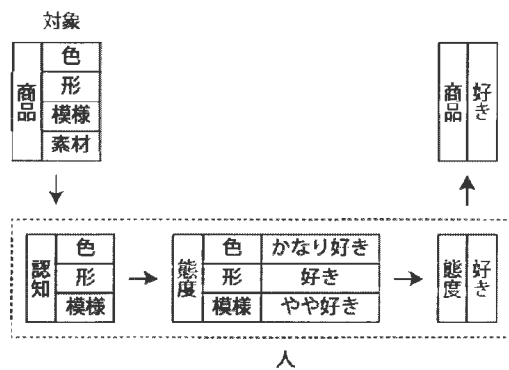


図1 商品に対する好みの態度を決定する過程

1. 商品の各属性を認知する。

ユーザは、各自のもつ動機(例:青いTシャツがほしい)にマッチする商品(例:Tシャツ)を見て、物の中から認知する。また、認知した商品においても、各自の動機に関連のある商品の属性(例:青)を注目して認知する。

2. 商品の各属性について好みの態度を決める。

認知した商品の各属性(例:Tシャツの色や首のデザインなど)について、各自の好みの基準により好みの態度をそれぞれ決定する。例えば、「Tシャツの色は好きだけど、首のデザインは、好きではない。」のように態度を決定する。

3. 各属性に関する好みの態度を連合して、その商品に関する最終的な態度を決定する。

認識した商品の各属性に関する好みの態度を連合させ、総合的に判断して商品に対する態度を決定する。例えば、あるTシャツについて、「色がとても気に入っている」と「胸に施されているロゴのデザインも悪くはない」という2つ属性に関する好みの態度から、この商品は好みであるという態度を決定する。

また本研究では、各ユーザの好みの基準は、そのユーザの重要視する属性に関連していると考えている。

例えば、商品A(色が青で、丸首のTシャツ)と商品B(色が青で、V首のTシャツ)に関する好みの態度の決定について考える。色に関心の高いユーザは、両商品に対して好意的な態度を示す可能性が高いと考えられる。一方、首のデザインに関心の高いユーザは、好みに適する商品にしか好意的な態度を示さないと考えられる。このようにユーザは、好みの基準として商品の属性の重要度を保持しており、その重要度がユーザの最終的な好みの態度に影響を与えていると考えられる。

本研究では、上記に示した、「ある商品に関する人の好みの態度の決め方」と「好みの基準」をベースとして、以下の3つのステップにより好みのモデル化を行う。

1. 商品属性の認知処理

ユーザが商品に対する好みの態度を決定するために不可欠である商品の属性・属性値をラフ集合[1,2]の縮約を用いて抽出する。ラフ集合の縮約とは、ある分類(例:好き)に属する対象を確実に他の分類(例:嫌い)の対象から識別するのに十分な極小の属性の集合のことである。

2. 属性(属性値)に関する好みの決定処理

ラフ集合の縮約により抽出した好みに関連する属性とその属性値に関するユーザの好み(好き・嫌い)の態度を度合いとして算出する。本研究では、この度合いを重要度と呼ぶ。属性に関する重要度は、商品の属性の内、どの属性に注目しているのかを指し示す。また属性値に関する重要度は、特定の属性において重要視している属性値を指し示す。

3. 商品に関する好みの決定処理

属性に重要度をユーザの好みの基準と捉え、商品の特徴を示す属性値に重みとして掛け合わせた度合いを嗜好度として算出する。本研究では、この嗜好度の集合を嗜好モデルと呼ぶ。この嗜好モデルは、ユーザが商品の特徴の内、どの属性値をどの程度好んでいるのかを表す。

以下の節において、嗜好モデルを算出する各ステップにおける情報処理の仕方について述べる。

2. 1 商品属性の認知処理

本研究では、ユーザが商品に関する好みの態度を決定するときに注視する属性・属性値を注視する過程をラフ集合によりモデル化する。

ラフ集合では、対象を識別する条件となる対象の属性(条件属性)の集合と、識別の目的となる属性(決定属性)の集合により表現した表を決定表と呼ぶ。また、決定属性の属性値に基づいて分類される対象の集合を決定クラスと呼ぶ。

表 1 商品の属性・属性値と商品に関する個人 X の好みの関係を示した決定表

対象	条件属性						決定属性
	Color (Co)	Texture (T)	Category (Ca)	Neck (N)	Sleeve (S)	Chest (Ch)	Preference
ID_1	Dark-Brown (DBr)	Point (Po)	Hood (Ho)	Crew (Cr)	Long (L)	Zip-up (Z)	Positive (P)
ID_2	Blue (Blu)	Point (Po)	Hood (Ho)	Crew (Cr)	Long (L)	Line (Li)	Positive (P)
ID_3	Black (Bla)	Uniform (U)	Cut Sewn (CS)	Henley (He)	Long (L)	Non (No)	Positive (P)
ID_4	Deep-Green (DGr)	Uniform (U)	Cut Sewn (CS)	Crew (Cr)	Long (L)	Non (No)	Negative (N)
ID_5	Pink (Pi)	Point (Po)	Hood (Ho)	Crew (Cr)	Long (L)	Mark (M)	Negative (N)
ID_6	Deep-Green (DGr)	Uniform (U)	Shirts (S)	Collar (Col)	Long (L)	Button (Bu)	Negative (N)

表 2 決定クラス D_p (好き) に関する決定行列

	ID_1	ID_5	ID_6
ID_1	{(Co, DBr), (T, Po), (Ca, Ho), (Ch, Z)}	{(Co, DBr), (Ch, Z)}	{(Co, DBr), (T, Po), (Ca, Ho), (Ne, Cr), (Ch, Z)}
ID_2	{(Co, Blu), (T, Po), (Ca, Ho), (Ch, Li)}	{(Co, Blu), (Ch, Li)}	{(Co, Blu), (T, Po), (Ca, Ho), (Ne, Cr), (Ch, Li)}
ID_3	{(Co, Bla), (Ne, He)}	{(Co, Bla), (T, U), (Ca, CS), (Ne, He), (Ch, No)}	{(Co, Bla), (Ca, CS), (Ne, He), (Ch, No)}

表 1 は、洋服を商品とし、洋服の色(Color)などの 6 つの条件属性とその属性値、決定属性の個人 P の好み(Preference)に関する評価とその値を示した決定表の例である。

表 1 の商品 $ID_1 \sim ID_6$ は、決定属性の属性値(好き(Positive)と嫌い(Negative))により 2 つの決定クラス D_p (好き) と D_n (嫌い) に分類される。

本論文では、ある分類に帰属する対象を確実に含めた確実性集合を表現する上で最小限必要な属性の集合を縮約と呼ぶ。本研究では、決定表から決定クラス D_p と D_n に属する商品を確実に類別する縮約を抽出する。例えば、 $ID_1 \in D_p$ と $ID_5 \in D_n$ は、属性 Color, または属性 Chest により確実に区別できる。このとき、属性 Color, または属性 Chest を、決定クラス D_p を識別するための縮約と呼ぶ。

この縮約を用いると、決定クラス D_k を識別する最小限の属性とその属性値の組み合わせを抽出できる。その組み合わせは、決定クラス D_k を導き出す IF-THEN ルールとして捉えることができる。

上記の例では、決定クラス D_p を導き出す属性とその属性値は、「IF Color=Dark Brown THEN D_p 」という IF-THEN ルールで記述できる。ラフ集合では、決定クラス D_k を識別する最小限の属性と属性値の組み合わせを極小決定ルールと呼ぶ。

縮約に基づく決定表内の全ての対象を確実に好きと嫌いに識別する極小決定ルールは複数存在する。そのため本研究では、決定行列を用いる方法 [3] を用いて、全ての極小決定ルールを抽出する。ここで、決定クラス D_p の決定行列を $M(D_p)$ とし、その成分 (i, j) を $M_{ij}(D_p)$ とする。決定行列 $M(D_k)$ から抽出する決定クラス D_k を識別する極小決定ルールの集合を $F(D_k)$ とする。表 2 は表 1 から作成した決定クラス D_p (好き) に関する決定行列 $M(D_p)$ である。

$M(D_p)$ の各成分は、 D_p (好き) を D_n (嫌い) と識別するために最小限必要な属性・属性値の組み合わせを示している。

表 3 好き (Positive) という態度に関する属性・属性値の決定ルール： $F(D_p)$

$F_i(D_p)$	属性=属性値	$F_i(D_p)$	属性=属性値
$F_1(D_p)$	Co=DBr	$F_4(D_p)$	Ch=Z
$F_2(D_p)$	Co=Blu	$F_5(D_p)$	Ch=Li
$F_3(D_p)$	Co=Blu	$F_6(D_p)$	Ne=He

表 4 好き (Positive) に関連する属性 (a)・属性値 (v_a) の重要度

a	重要度： $ID_{D_p}(a)$	重要度の比率
Ne	0.33	0.16
Co	1.00	0.50
Ch	0.66	0.33

v_a	重要度： $ID_{D_p}(v_a)$	重要度の比率
Bl	0.33	0.16
He	0.33	0.16
Blu	0.33	0.16
DBr	0.33	0.16
Z	0.33	0.16
L	0.33	0.16

表 5 洋服に関する個人 X の嗜好度

属性値	嗜好度	属性値	嗜好度
Blu	0.27	No	-0.01
DBr	0.27	U	-0.02
Z	0.23	Cr	-0.04
Li	0.22	Col	-0.13
He	0.13	M	-0.23
CS	0.00	Bu	-0.23
Sh	0.00	Pi	-0.27
		DGr	-0.27

*注：値は小数点 3 位で四捨五入

例えば、表 2 に示した決定行列 $M(D_p)$ の成分 $M_{12}(D_p) : \{(Co, DBr), (Ch, Z)\}$ は、 $ID1 \in D_p$ と $ID5 \in D_n$ とを属性:Color(Co)の属性値:Dark-Brown(DBr), または属性:Chest(Ch)の属性値:Zip-up(Z)により識別できることを示している。

決定行列 $M(D_p)$ に内在する極小決定ルール $F(D_p)$ は、以下の通りとなる。

$$F(D_p) = ((Co, DBr) \vee (Ch, Z)) \vee ((Co, Blu) \vee (Ch, Li)) \vee ((Co, Bla) \vee (Ne, He))$$

上記の論理式より、好き (Positive) という態度を決定するために不可欠な属性・属性値の組み合わせは、表 3 まとめられる。同様に嫌い (Negative) という態度を決定するために不可欠な属性・属性値について求めることができる。

2.2. 属性 (属性値) に関する好みの決定処理

2.1 節で極小決定ルールとして抽出した属性とその属性値について、重要度を算出する。

ここで、極小決定ルールに含まれる属性を a としたとき、その取りうる属性値 v_a とする。そのとき、決定クラス (D_k に属する全ての商品 $O(D_k)$ の内において、決定ルール $F_i(D_k)$ を含む商品の割合 (被覆度) : $CI(F_i(D_k))$ を式 (1) とする。

式 (1) :

$$CI(F_i(D_k)) = F_i(D_k) \cap O(D_k) / O(D_k)$$

例えば、表 3 に示した決定クラス D_p (好き) の極小決定ルール $F_1(D_p) = [Co = DBr]$ の被覆度は、 $CI(F_1(D_p)) = 1/3$ となる。

また、極小決定ルールを構成する属性・属性値の個数 (ルールの長さ) を $Len(F_i(D_k))$ とする。ここで、決定ルールの集合 $F(D_k)$ の内で、属性 a を含む極小決定ルールの数を n とすると、決定クラス D_k の属性 a の重要度 : $ID_{D_k}(a)$ を式 (2) と定義する。

式 (2) :

$$ID_{D_k}(a) = \sum_{i=1}^n \left(\frac{CI(F_i(D_k))}{Len(F_i(D_k))} \times \frac{1}{(Len(F_i(D_k)))^e} \right)$$

ただし、ここで用いる $CI(F_i(D_k))$ は、属性 a を含んだ $F_i(D_k)$ に関する $CI(F_i(D_k))$ の値とする。

同様に、 $F(D_k)$ の内で属性値 v_a を含む極小決定ルールの数を m とすると、決定クラス D_k の属性値 v_a の重要度 : $ID_{D_k}(v_a)$ を式 (3) と定義する。

式 (3) :

$$ID_{D_k}(v_a) = \sum_{j=1}^m \left(\frac{CI(F_j(D_k))}{Len(F_j(D_k))} \times \frac{1}{(Len(F_j(D_k)))^e} \right)$$

ただし、ここで用いる $CI(F_i(D_k))$ は、属性値を含んだ $F_i(D_k)$ に関する $CI(F_i(D_k))$ の値とする。

例えば、表 3 より決定クラス D_p (好き) の属性 Color の重要度は、 $n = 3$, $CI(F_1(D_p)) = CI(F_2(D_p)) = CI(F_3(D_p)) = 1/3$, $Len(F_1(D_p)) = Len(F_2(D_p)) =$

$Len(F_3(D_p)) = 1$ より、以下の通りとなる。

$$ID_{D_p}(color) = \left(\frac{1}{3} \times \frac{1}{1^e}\right) + \left(\frac{1}{3} \times \frac{1}{1^e}\right) + \left(\frac{1}{3} \times \frac{1}{1^e}\right) = 1.0$$

決定クラス D_p (好き)の属性Chestと属性Neckの重要度を計算すると、それぞれ $ID_{D_p}(chest) = 0.66$ 、 $ID_{D_p}(neck) = 0.33$ となる。3つの属性の重要度を比較することで、個人Xにおいては属性Colorが最も好きに関連している属性であることがわかる。表4は、好きに関する極小決定ルールに出現する属性と属性値の重要度とその比率を示した表である。決定クラス D_n (嫌い)についても、 D_p (好き)と同様に各属性・属性値に関する重要度を算出する。

各属性・属性値に関して好きと嫌いの重要度を統合し、総合的に属性・属性値の好みへの重要度として、統合重要度を算出する。式(4)と(5)に、属性の統合重要度： $UD(a_i)$ と属性値の統合重要度： $UD(v_{a_i})$ の算出式を示す。

式(4)：

$UD(a_i)$

$$= \begin{cases} ID_P(a_i) \times ID_N(a_i), & ID_P(a_i), ID_N(a_i) \neq 0 \\ 0 \times ID_N(a_i), & ID_P(a_i) = 0 \\ ID_P(a_i) \times 0, & ID_N(a_i) = 0 \end{cases}$$

式(5)：

$UD(v_{a_i})$

$$= \begin{cases} ID_P(v_{a_i}) - ID_N(v_{a_i}), & ID_P(v_{a_i}), ID_N(v_{a_i}) \neq 0 \\ 0 - ID_N(v_{a_i}), & ID_P(v_{a_i}) = 0 \\ ID_P(v_{a_i}) - 0, & ID_N(v_{a_i}) = 0 \end{cases}$$

好きと嫌いの両方に関して重要度の高い属性は、個人Xにより注視されおり、商品の好きと嫌いを評価に寄与していると捉えることができる。そのため本研究では、属性に関する統合重要度を好き・嫌いの属性の重要度を乗算することで算出する。これにより、個人の注目している属性の重要度はより強調し、注目されていない属性の重要度を低下させる。

また、好きと嫌いの両者に属性値が現れるということは、その個人にとってその属性値に対する評価は定まっていないことを表している。したがって本研究では、属性値に関する統合重要度を好きから嫌いの属性値の重要度を減算することで算出する嫌いにのみ出現する属性値の重要度や両者に出現する属性の重要

度も低下させる。

2.3. 興味度の算出

好きと嫌いに関する属性・属性値の統合重要度から、個人Xの嗜好モデル $PM(X)$ を作成する。嗜好モデル $PM(X)$ の成分を $PM_i(X)$ とすると、 $PM_i(X)$ は以下のように定義できる。

表6 洋服を表現する属性とその属性値

属性	属性値
Color (Co)	Black, Blue, Brown, Dark-Brown, Deep-Blue, Deep-Green, Gray, Green, Pink, Purple, Red, Red-Purple, Yellow
Texture (T)	Check, Horizontal-Stripes, Point, Uniform
Category (Ca)	Cut-Sewn, Hood Polo-Shirts, Shirts, Knit
Neck (N)	Collar, Crew, Henley, High, Hood Turtle, V
Sleeve (S)	Half, Long
Chest (Ch)	Button, Mark, Zip-Up, None

表7 各セットに対する解答属性値

セット	好きと判断する基準
1	属性値=色：青 (Blue, Deep-Blue)
2	属性値=首：ヘンリーネック (Henley)

表8 評価実験データより算出した嗜好度

セット番号 (解答属性値)	各属性値の嗜好度 (上位5位)	
データセット：1 (Blue と Deep-Blue)	Deep-Blue	+0.14
	Blue	+0.11
	Gray	+0.04
	Uniform	+0.01
	Half	+0.01
データセット：2 (Henley)	Henley	+0.26
	Gray	+0.10
	Black	±0.00
	Uniform	±0.00
	None	±0.00

$PM_i(X)$ を個人Xの属性値*i*に関する嗜好度と呼ぶ。式(6)：

$$PM_i(X) = \{UD(v_{a_i}) \times UD(a_i)\}$$

本研究では、式(6)に基づいて算出される嗜好度

$PM_i(X)$ の値が正 (+) 値の場合、その個人の好きに起因している属性値と考え、負 (-) 値は嫌いに起因している属性値と捉える。表 5 は、個人 X の嗜好モデルの各成分である嗜好度 $PM_i(X)$ を降順にまとめた例である。嗜好度 $PM_i(X)$ の値より対象の色が青や黒の場合、個人 X は好むと推定できる。

3. 嗜好モデルの評価

洋服 Z に関する個人 X の好みの態度が「青色」に関係している場合、青色の嗜好度の値が正 (+) 値で、かつ、他の属性値の嗜好度より相対的に高い値を示したならば、嗜好モデルは個人 X の嗜好を推定できていると見なせる。

上記の考えに基づいて、被験者に対象を好きと判断する基準とする属性値を与えた上で、洋服の好みを判断してもらい、嗜好モデルにより好みの判断基準とした属性値の推定する評価実験を行った。本評価実験では、嗜好モデルの評価を以下の手順で行った。

1. 男性用上着 520 着からランダムに選択した 50 着を 1 セットとし、2 セットの決定表データを用意した。表 6 は、本実験において対象を表現するのに用いた属性と属性値である。
2. 対象を好きと判断する基準とする属性値（以後、解答属性値）を被験者に与える。一人の被験者に対して、2 セットの実験を行った。表 7 は、各セットに対して設定した解答属性値を示している。
3. 被験者は、解答属性値を含む洋服を全て好きと判断する。それ以外の洋服は、被験者の判断に任せる。例えばデータセット 1 の場合、解答属性値である青 (Blue と Deep-Blue) を含む対象は、無条件で好きと答え、それ以外は被験者の判断に任せる。
4. 評価データから嗜好度を算出する。解答属性値とした属性値の嗜好度が正 (+) 値であり、かつ、全属性値における相対的な割合が他の属性値と比べて高ければ、嗜好モデルは好みを推定できたとする。

表 8 は、被験者 X がデータセット 1 と 2 に対して行った評価実験データから嗜好度を算出した結果を示している。

データセット 1 に関しては、解答属性値に設定した Blue と Deep-Blue の嗜好度の値が 1, 2 位となっている。データセット 2 も同様に、解答属性値に設定した Henley の嗜好度の値が 1 位となっている。この結果から、本論文で提案した嗜好度により人の好みを推定できていることが示された。

4. まとめ

本論文では、個人適応型の情報検索システムを実現するために、商品に関して人が好みを決定する過程を以下の 3 つの過程と捉え、その過程をシステム上で模倣するための嗜好モデルを提案した。

1. 商品の各属性を認知する。
2. 商品の各属性について好みの態度を決める。
3. 各属性に関する好みの態度を連合して、その商品に関する最終的な態度を決定する。

嗜好モデルを作成する方法として、ラフ集合の縮約を用いて対象に関する個人の好みに関連する対象の属性・属性値を抽出し、かつ、対象の属性・属性値と対象の関する好みの関連の度合い（嗜好度）として算出した。人が対象の好みを判断するときに重要視している属性値を嗜好モデルにより推定できるかを主観評価実験のデータから嗜好モデルを生成して評価を行った。その結果、主観評価実験において被験者が好みの判断を行う上で重要視していた対象の属性値を、本論文で提案した嗜好モデルにより推定できた。

今後の課題は、この嗜好モデルを情報検索システムに適用し、情報検索システムが利用者に提示した結果の理由が利用者に直感的にわかる情報検索システムを開発する。

参考文献

- 1) 森典彦, 田中英夫, 井上勝雄: “ラフ集合と感性—データからの知識獲得と推論”, 海文堂出版, 2004
- 2) Pawlak, Z.: “Rough Sets”, *International Journal of Computer & Information Sciences*, vol.11, issue 5, pp.341-356, 1982
- 3) Shan, N., Ziarko, W.: “Data-based acquisition and incremental modification of classification rules”, *Computational Intelligence*, vol.11, pp.357-370, 1995.