

SVMを用いたファッション推薦システムについて

吉田, 慎一郎
九州大学

馬場, 隆寛
九州大学

村上, 直至
九州大学

清水, 一憲
九州大学

他

<http://hdl.handle.net/2324/1518762>

出版情報 : 火の国情報シンポジウム論文集. 2011, pp.1-8, 2011-03. 情報処理学会
バージョン :
権利関係 : (C) 2011 Information Processing Society of Japan



SVM を用いたファッション推薦システムについて

吉田慎一郎^{†1} 馬場隆寛^{†1}
村上直至^{†1} 清水一憲^{†1}
殷成久^{†1} 伊東栄典^{†1} 廣川佐千男^{†1}

感性的な要素が多く含むファッション情報は、単純なテキスト解析では困難な課題である。本発表では、アンケートからユーザープロフィールを抽出し、SVM を用いてユーザーの嗜好を推定するシステムを提案し、その実験結果について報告する

Using Support Vector Machines for Fashion Recommendation System

Shinichiro Yoshida^{†1} and Takahiro Baba^{†1}
Naomichi Murakami^{†1} and Kazunori Shimizu^{†1}
Chengjiu Yin^{†1} and Eisuke Itoh^{†1} and Sachio Hirokawa^{†1}

It is difficult to analyze fashion information through the analysis of text collection because fashion is something that is influenced by people's emotions and feelings. In this paper, we extract the user's profiles from the questionnaire, and then propose an emotion presumption system to predict user's preferences. In this system, a Support Vector Machine is used to analyze fashion information.

1. はじめに

情報検索のための技術的な研究は、インターネットの普及を契機として、質・量ともに飛躍的に進歩しており、インターネットによる情報収集が可能になっている。何か調べ事をするときには、我々は検索エンジンを使う[1]。検索の存在なしでは、インターネットの使い勝手は半減してしまう、といっても過言でもない。

しかし、現在の検索エンジンはキーワード検索の方式である。検索に慣れていない人にとっては、キーワード検索エンジンを使いこなすことは簡単ではない。そこで、大量のデータを分析し、利用者の特徴、性格、傾向、興味などを見つけ出し、それらを情報の精度を向上させるための次世代の検索エンジンが必要である。

次世代の検索エンジンはデータマイニング技術を活用した検索エンジンである。データマイニングの技術は大量のデータを自動的に分析して、何らかの意味のある有用な傾向・パターン、規則・ルール、原理・原則などを抽出する。検索・分析だけでなく、好み・傾向・ニーズに合わせて適切なものを推薦することも必要となる。次世代の検索エンジンを構築するために、以下のようなことが必要となる。

- a) インターネット上のデータ収集
- b) キーワード検索エンジンの作成
- c) データの分析
- d) 推薦

筆者の所属する研究室でも、a)、b)、c)に関する研究を行ってきている [2][3][4]。本研究は研究対象をファッションのデータとして「c)データの分析」を行う。

本研究はアパレル業界に着目した。その理由としては分析対象のデータが入手できたことが大きい。しかし、現在のアパレル業界ではマーケットの変化、個人消費低迷、ニーズの多様化、ネット・通販の躍進などの問題が起きている。このような時代に、ユーザー傾向などを分析し、その好みに合わせて、商品を提供しないといけない。そこで、本研究はファッションデータを対象とし、サポートベクターマシン(以下 SVM)による分析手法を提案する。

本研究は通常の因子分析とは違う。本研究では、アンケート情報を使い、被験者の一対比較の選好結果を予測することを考え、交差検定を使用し、推定しやすい・しにくい商品対があるか否かを分析する。また、被験者の嗜好推測の手がかりとなるようなアンケート項目を特定する。

2. 分析対象データ

昨年(2010年)、日本オペレーションズ・リサーチ学会実践的データマイニング研

^{†1}九州大学
Kyushu University

究部会が主催し、人工知能学会、経営情報学会関西支部、情報処理学会数理モデル化と問題解決研究会、株式会社玉屋、チームラボ株式会社、株式会社数理システム、ERATO 湊離散構造処理系プロジェクト、ビジネスマイニング研究センター有限責任事業組合が協賛する第 2 回リコメンデーションコンテストが実施されている[5]。

実践的データマイニング研究部会は「選好実験システム」を使って、619 名の被験者(女性)に対して、アンケートを行った。図 1 にアンケート収集のおおまかな流れを示す。

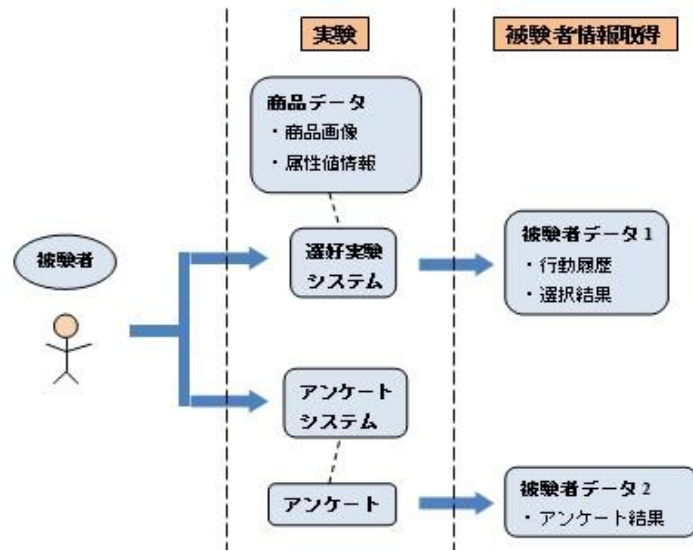


図 1. アンケート収集の流れ

選好実験システムでは、最初に実験に関する説明が表示される。その後、商品対画像、および属性(ブランド名、価格、素材、セールスポイント)を隠したものの(クリックすればみることができる)が表示され、被験者は左右のどちらかの商品を選択する。その作業を 28 の商品対すべてに対して行う。

アンケートシステムでは、ファッションに関する質問が 14 個表示され、被験者はそのアンケートに回答する。この実験により、被験者の画像選択に関する情報(属性確認の有無や選択にかかった時間等の行動履歴、選択結果等)、およびファッションに関するアンケートの回答結果が得られる。日本オペレーションズリサーチ学会実践的データマイニング研究部会が主催する第 2 回リコメンデーションコンテストで提供されたデータの中で、本論文で使用したデータを表 1、表 2、表 3 に示す。

表 1. ファッション選択データ

選好実験システム	商品データ	ア)画像データ.docx	実験で使用した商品の画像データ(訓練データの画像のみ使用)
	被験者データ 1	イ)train.csv	被験者が選択した訓練用の実験データに対する結果

表 2. アンケートデータ

アンケートシステム	アンケート	ウ)アンケート項目	アンケート項目一覧データ
	被験者データ 2	エ)quest_unify.csv	被験者に対するアンケートに関する結果

表 3. その他のデータ

その他	オ)アンケート調査の取得例.swf	アンケートを行った実際のサイトでの取得方法
-----	-------------------	-----------------------

3. SVM

代表的な機械学習の手法として、SVM, SVR, SMO, 決定木, k 近傍法などがある。SVM は二値分類器として精度が高い。また、本論文で扱う一対比較問題に適用できると考えたため SVM を使用した。

3.1 SVM 例

被験者 6 人に対して、ある左右に並んだ商品対が提示され、また、表 4 のようなアンケートを行い、表 5 のような 6 人の選好結果およびファッションに関するアンケート結果が得られたとする。表 5 を SVM_Light を使用するためのデータ形式に変換したものを表 6 に示す。ラベルは、L を選択した場合は 1、R を選択した場合は -1 である。特徴ベクトルの ID は、問 1 の A,B、問 2 の A,B、問 3 の A,B,C,D をそれぞれ 1,2,3,4,5,6,7,8 とし、要素は、選択しているものは 1、選択していないものは 0 とする。

表 6 に SVM_Light を適用して求めた特徴ベクトルの重みを表 7 に示す。表 7 を見ると、5(カジュアル)、7(かわいらしさ)は L を、6(上品)、8(スタイリッシュ)は R を選択する要因として影響を与えている。特に 7(かわいらしさ)は他のものに比べてより大きな影響を与えている。この結果は、6 人のデータに対してはいえるが、未知のものに対していえるかどうかはわからない。

表 4. アンケート

問 1	店員のアドバイスを参考にしますか？	A はい B いいえ
問 2	着心地とファッション性のどちらを優先しますか？	A 着心地 B ファッション性
問 3	ファッションで重視するイメージは？	A カジュアル B 上品 C かわいらしさ D スタイリッシュ

表 5. 選好結果の例 (L,R は左側、右側の商品を選択したという意味である)

名前	選択	問 1		問 2		問 3			
		A	B	A	B	A	B	C	D
山田	L	○			○	○		○	
鈴木	L	○		○				○	
吉田	L		○	○				○	
木村	R	○		○			○		
田中	R	○			○		○		
佐藤	R		○	○					○

表 6. SVM_Light 使用のためのデータ形式化

ラベル L:1,R:-1	特徴ベクトル(ID:要素)								
1	1:1	2:0	3:0	4:1	5:1	6:0	7:1	8:0	
1	1:1	2:0	3:1	4:0	5:0	6:0	7:1	8:0	
1	1:0	2:1	3:1	4:0	5:0	6:0	7:1	8:0	
-1	1:1	2:0	3:1	4:0	5:0	6:1	7:0	8:0	
-1	1:1	2:0	3:0	4:1	5:0	6:1	7:0	8:0	
-1	1:0	2:1	3:1	4:0	5:0	6:0	7:0	8:1	

表 7. 特徴ベクトルの重み

特徴ベクトル ID								
1	2	3	4	5 カジュアル	6 上品	7 可愛らしさ	8 スタイリッシュ	
0	0	0	0	0.32	-0.63	0.95	-0.32	

3.2 本研究の特徴

通常、因子分析の結果はそのデータに対しては最適な関係であるが、その他のデータに対して適しているかどうかは分からない。

推薦システムではこのような未知のデータに対して正答率を評価しなければならない。通常、推薦システムの評価では交差検定が用いられる。

交差検定の種類の中に K-分割交差検定というものがある。標本群を K 個に分割し、1 つをテストデータ、K-1 個を訓練データとする。K 個に分割された標本群それぞれをテストデータとして K 回検定を行い、その結果を平均して 1 つの推定を得る。

本論文では、10 分割交差検定を使用し、推定しやすい・しにくい商品、アンケート項目の有効性について分析する。

4. 画像対の提示によるユーザーの好みの測定

4.1 測定結果

まず、「ア」実験で使用した商品の画像データの「訓練データ」の商品対について、上から商品対 1,2,3,・・・,21 と呼ぶことにし、その 1,2,3,・・・,21 を商品対番号と呼ぶことにする。また、各被験者には ID がつけられており、これをソートし順番に被験者 1,2,3,・・・,619 と呼ぶことにし、その 1,2,3,・・・,619 を被験者番号と呼ぶことにする。

21 個の商品対(訓練用の実験データ)を使って、619 名の被験者(女性)の選択を集計した。その中から特徴的なものを 2 つピックアップし、それぞれ商品画像とともに図 2、3 に示す。(L,R はそれぞれ左側、右側の商品を意味し、数字はそれを選択した人数である。)

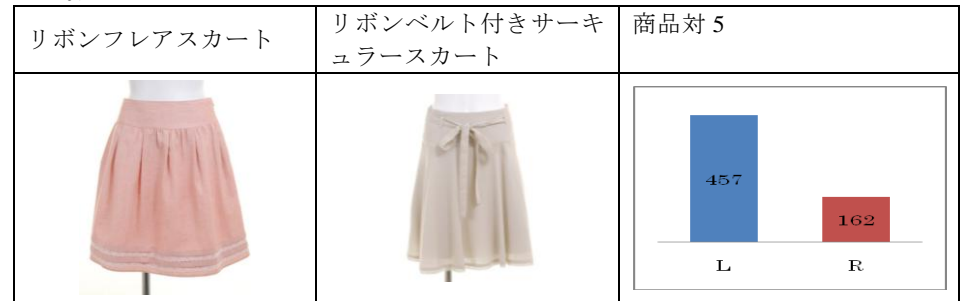


図 2 商品対 5 の商品画像と選択結果



図 3 商品対 11 の商品画像と選択結果

図 2 は選択に偏りがある商品対である。商品対の画像を見ると、形に大きく違いがあり、色も異なっている。このように偏りがある商品対は、どちらを選択するか予測がしやすいのではないだろうか。

図 3 は選択にあまり偏りがない商品対である。商品対の画像を見ると、形に大きな違いはなく、色が異なっている。このようにあまり偏りがない商品対は、どちらを選

択するか予測がしにくいのではないだろうか。

このような分析から、筆者は、予測の難易度と好みの偏りに何らかの関連があるのではないかと考えるようになった。本稿では、予測の難易度を SVM による推定の予測的中率として定式化することにした。

4.2 分析

推定しやすい・しにくい商品対とはどういう特徴を持つのかを、SVM による推定値、予測的中率、画像選択データの観点から調べる。

4.2.1 データセット

用いたデータセットは、21 の各商品対についてそれぞれ 619 サンプル(被験者 619 名分)で、ラベルは、被験者が左の画像を選んだ場合は 1(正例)、右の画像を選んだ場合は-1(負例)とする。

また、14 個の質問の選択肢(以降、選択肢項目と呼ぶ)をすべて独立と考え、質問 1~14、選択肢項目の左から右へ選択肢番号を 1 から順番に振る。なお、選択肢が多数となる質問 13 は表 8 のように、著者が 5 段階に分けたものを使用した。選択肢項目数は 93 となり、選択肢番号を特徴ベクトル ID とし、特徴ベクトルの要素値は選択肢項目において、選択されたものを 1、選択されていないものを 0 とした。

表 8 年齢データの取り扱い

質問 13.あなたの年齢は？				
0.18 歳以下	1.19~24 歳	2.25~29 歳	3.30~34 歳	4.35 歳以上

4.2.2 偏りと SVM による的中率の関係

SP 率、PB 率、CB 率

4.2.1 のデータセットを用いて SVM_Light を使用し 10 分割交差検定を行い、各商品対について各評価データに対する推定値を 619 名分取得した。

商品対 p において、被験者 i の推定値が 0 以上である場合、その被験者は左の商品を選択すると予測したこととし、

$$P_L(p, i) = 1, P_R(p, i) = 0$$

とする。推定値が 0 よりも小さい場合は右の商品を選択すると予測したこととし、

$$P_L(p, i) = 0, P_R(p, i) = 1$$

とする。

また、商品対 p において、被験者 i が実際に左の商品を選択していた場合、

$$C_L(p, i) = 1, C_R(p, i) = 0$$

とし、右の商品を選択していた場合は、

$$C_L(p, i) = 0, C_R(p, i) = 1$$

とする。

表 9. 予測結果と選択結果の一致者の合計人数(単位は人)

商品対番号	$L(p)$	$R(p)$
1	424	0
2	192	162
3	323	20
4	61	324
5	456	0
6	442	0
7	0	429
8	225	134
9	329	38
10	95	284
11	230	125
12	194	127
13	395	23
14	20	346
15	428	0
16	313	44
17	501	0
18	384	21
19	374	2
20	8	419
21	126	253

ここで、サンプル数を $n(=619)$ とし、商品対 p について、次の $L(p)$ 、 $R(p)$ 、および $M(p)$ を定義する。

$$L(p) = \sum_{i=1}^n P_L(p, i) \cdot C_L(p, i)$$

$$R(p) = \sum_{i=1}^n P_R(p, i) \cdot C_R(p, i)$$

$$M(p) = X(p) + Y(p)$$

$L(p)$ 、 $R(p)$ はそれぞれ、商品対 p において、左を選択すると予測されたかつ実際に左を選択した被験者数の合計、右を選択すると予測されたかつ実際に右を選択した被験者数の合計である。 $M(p)$ は $L(p)$ と $R(p)$ の合計である。

全ての商品対の $L(p)$ および $R(p)$ を表 9 に示す。そのとき、商品対 p について、次の SP 率 (Success Predict Rate: 予測的中率) を定義する。

$$SP(p) = \frac{M(p)}{n} \times 100$$

例として、 $p=8$ (商品対 8) の場合について計算してみる。表 9 より、 $L(8)=225$ 、 $R(8)=134$ なので、

$$SP(8) = \frac{M(8)}{n} \times 100 = \frac{L(8) + R(8)}{n} \times 100 = \frac{225 + 134}{619} \times 100 = 58.00$$

となる。また、商品対 p について、次の $Pl(p)$ 、 $Cl(p)$ を定義する。

$$Pl(p) = \begin{cases} \sum_{i=1}^n P_L(p, i), & \text{if } \sum_{i=1}^n P_L(p, i) > \sum_{i=1}^n P_R(p, i) \\ \sum_{i=1}^n P_R(p, i), & \text{if } \sum_{i=1}^n P_L(p, i) < \sum_{i=1}^n P_R(p, i) \end{cases}$$

$$Cl(p) = \begin{cases} \sum_{i=1}^n C_L(p, i), & \text{if } \sum_{i=1}^n C_L(p, i) > \sum_{i=1}^n C_R(p, i) \\ \sum_{i=1}^n C_R(p, i), & \text{if } \sum_{i=1}^n C_L(p, i) < \sum_{i=1}^n C_R(p, i) \end{cases}$$

表 10. 予測結果の人数と選択結果の人数 (単位は人)

商品対 p	SPL	SPR	SCL	SCR	$Pl(p)$	$Cl(p)$
1	619	0	424	195	619	424
2	327	292	322	297	327	322
3	557	62	365	254	557	365
4	117	502	239	380	502	380
5	618	1	457	162	618	457
6	619	0	442	177	619	442
7	0	619	190	429	619	429
8	374	245	336	283	374	336
9	534	85	376	243	534	376
10	177	442	253	366	442	366
11	390	229	334	285	390	334
12	359	260	327	292	359	327
13	582	37	409	210	582	409
14	57	562	236	383	562	383
15	619	0	428	191	610	428
16	525	94	363	256	525	363
17	619	0	501	118	619	501
18	584	35	398	221	584	398
19	613	6	378	241	613	378
20	12	607	196	423	607	423
21	220	399	274	345	399	345

$Pl(p)$ は左を選択すると予測された被験者数と右を選択すると予測された被験者数のうち多いほうを表し、 $Cl(p)$ は左を選択した被験者数と右を選択した被験者数のうち多いほうを表す。

各商品対において、 SPL と SPR 、 SCL と SCR 、および Pl と Cl の人数を表 10 に示す。

$$SPL = \sum_{i=1}^n P_L(p, i), SPR = \sum_{i=1}^n P_R(p, i), SCL = \sum_{i=1}^n C_L(p, i), SCR = \sum_{i=1}^n C_R(p, i)$$

そのとき、商品対 p について、PB 率(Predict Bias Rate: 予測偏り率)と CB 率(Choice Bias Rate: 選択偏り率)を定義する。

$$PB(p) = \frac{Pl(p)}{n} \times 100 \quad CB(p) = \frac{Cl(p)}{n} \times 100$$

例として、 $p=3$ (商品対 3)の場合の PB 率と、 $p=21$ (商品対 21)での CB 率を計算してみる。表 10 より、 $Pl(3)=557$ なので、

$$PB(3) = \frac{P_{lot}(3)}{n} \times 100 = \frac{557}{619} \times 100 = 89.98$$

となる。また、 $Cl(21)=345$ なので、

$$CB(21) = \frac{C_{lot}(21)}{n} \times 100 = \frac{345}{619} \times 100 = 55.74$$

となる。商品対 p における SP 率、PB 率、CB 率を表 11 に示す。

表 11. SP 率, PB 率, CB 率(単位は%)

商品対 p	$SP(p)$	$PB(p)$	$CB(p)$
1	68.50	100.00	68.50
2	57.19	52.83	52.02
3	55.41	89.98	58.97
4	62.20	81.10	61.39
5	73.67	99.84	73.83
6	71.41	100.00	71.41
7	69.31	100.00	69.31
8	58.00	60.42	54.28
9	59.29	86.27	60.74
10	61.23	71.41	59.13
11	57.35	63.00	53.96
12	51.86	58.00	52.83
13	67.53	94.02	66.07
14	59.13	90.79	61.87

15	69.14	100.00	69.14
16	57.67	84.81	58.64
17	80.94	100.00	80.94
18	65.43	94.35	64.30
19	60.74	99.03	61.07
20	68.98	98.06	68.34
21	61.23	64.46	55.74

SP 率、PB 率、CB 率の相関関係

表 11 を用い、エクセルにより算出した PB 率-SP 率、PB 率-CB 率、SP 率-CB 率の相関関係を図 4、図 5、図 6 に示す。図 4 を見ると、正の相関がある。ただ、予測偏り率が 85~100% の付近では大きくばらつきが見られる。図 5 を見ると、正の相関がある。これも図 4 と同じく、予測偏り率が 85~100% の付近では大きくばらつきが見られる。図 6 を見ると、強い正の相関がある。つまり、選択が片方に偏るような商品対は予測がかなり当たっている。一方、選択がきれいに分かれるような商品対については全然当たらない。

予測当たり率から推定しやすい・しにくい商品対を判断すると、選択が偏りそうな商品対が推定しやすい商品対、きれいに分かれそうな商品対が推定しにくい商品対ということになる。しかし、これらのことは人の目で見てわかりそうである。したがって、今回の方法では推定しやすい・しにくい商品対がどのような特徴を持つかわく説明はできない。

そこで次章以後では、予測的中率に影響を与える被験者属性がどのようなものであるかを考察する。

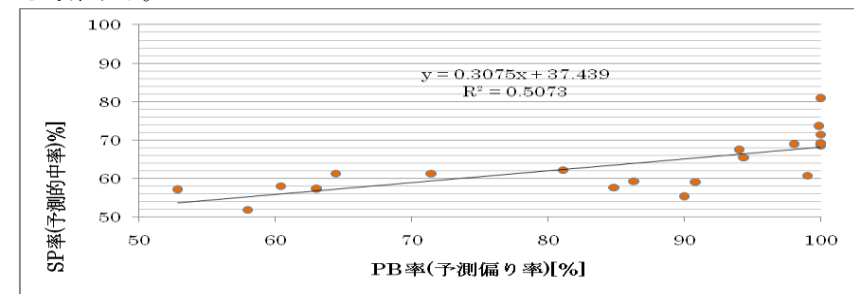


図 4. PB 率-SP 率

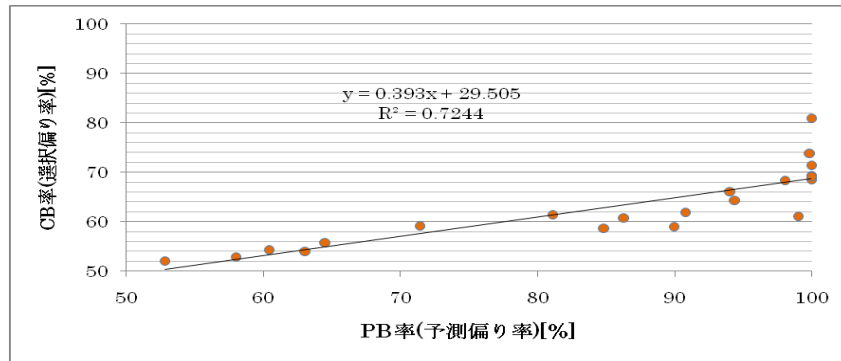


図 5. PB 率-CB 率

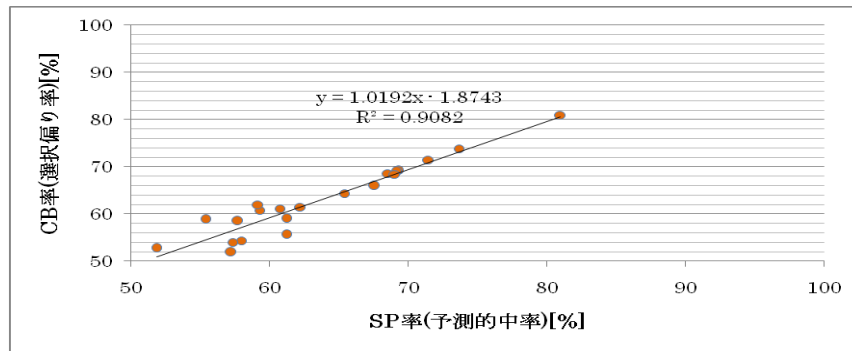


図 6. SP 率-CB 率

5. アンケート項目の有効性評価

5.1 選択肢項目の重み

4.2.1 のデータセットを用いて SVM_Light を使用し 10 分割交差検定を行い、各商品対について訓練ごとに特徴ベクトルに対する重みを求めた。各商品対について求めた訓練データごとの重みを平均した。

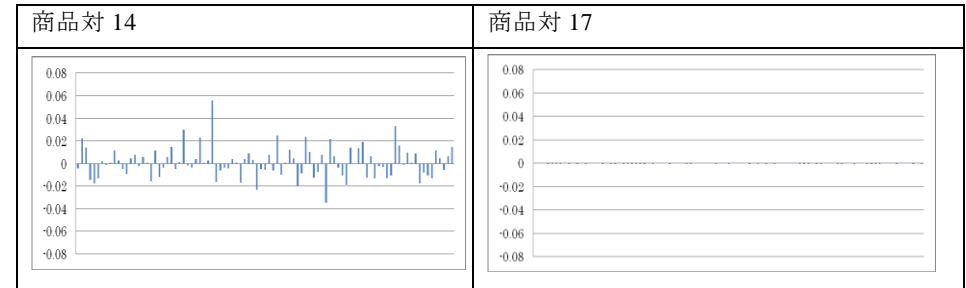
その中から特徴的なものを 2 つピックアップし、同スケールで表 12 に示す。なお、縦軸は重み、横軸は選択肢項目であり、一番左が選択肢番号 1、一番右が選択肢番号 93 である。

商品対 14 を見ると、いくつか重みの大きなものが見受けられる。この商品対において、これらの選択肢項目は他の選択肢項目に比べて、選好への影響が大きいと考え

られる。

商品対 17 を見ると、商品対 14 に比べて各選択肢項目の重みが非常に小さい。この商品対については、どの選択肢項目も選好への影響がほぼない。

表 12. 商品対 14,17 に対する選択肢の重み



5.2 影響度

選択肢番号 j の重みを w_j としたとき、影響度 $I(j)$ を次のように定義する。

$$I(j) = |w_j|$$

各商品対についての影響度を計算し、それらの上位 3 つを集計し質問ごとに分け、質問文のまとめとともに図 7 に示す。

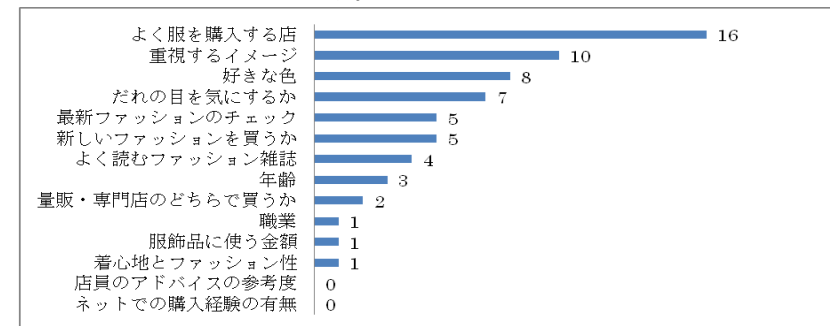


図 7. 各商品対における影響度上位 3 つの集計結果

5.3 分析

5.3.1 上位にある項目の分析

図 7 を見ると、上位には「よく服を購入する店」、「重視するイメージ」「好きな色」、「だれの目を気にするか」という内容の質問がある。

よく服を購入する店 一般に、店ごとにコンセプトなどが異なり取り扱う商品群も

違う。よく服を購入する店が異なると嗜好も異なりやすくなり影響が大きくなったと考えられる。

重視するイメージ これが大きいということは、当たり前のように商品画像と自分の重視するイメージとを照らし合わせて選好しているように思える。

好きな色 形やデザインがほぼ同じような商品対7や11の場合、形やデザインなどにより選ばれるかもしれないが、おそらく色で選ばれるだろう。また、実験では2つの商品のうち好きな方1つを選択しなければならないが、もしかしたらどちらも好きではない場合も考えられる。その場合は、嫌いな色よりも好きな色のものを選びそうである。よって影響が大きくなったと考えられる。

だれの目を気にするか この項目の影響が大きいということは、好みの商品を選択する際に自分がその商品を着たときのことを考えているのかもしれない。

5.3.2 下位にある項目の分析

下位には「ネットでの購入経験の有無」、「店員のアドバイスの参考度」、「着心地、ファッション性」、「服飾品に使う金額」、「職業」という内容の質問がある(図7)。

ネットでの購入経験の有無 今回の実験は「実際に購入するなら(ネットで)どちらを選ぶか」というものではなく「好きなのはどちらか」であるため影響が小さかったと考えられる。

店員のアドバイスの参考度 今回の実験は店員のアドバイスを聞きながら行っていただけではなく、被験者が自分で好きな方を選んだだけであるため影響が小さかったと考えられる。

着心地、ファッション性 ファッション性は画像を見ただけでもある程度分かるが、着心地は画像を見ただけでは分からない。今回の実験では属性として各商品の素材を見ることができたが、見たとしても「ポリエステル60%」などと書かれているだけなのでその商品の着心地は分からないだろう。よって影響が小さかったと考えられる。

服飾品に使う金額 今回の実験では属性として各商品の値段を見ることができたが、実際に買うわけではないので影響が小さかったと考えられる。

職業 提供されたデータを見ると、19~29歳の割合が約8割である。その幅の年齢の人たちの選好に「年齢」は大きな影響を与えないことがわかる。

6. まとめと今後の課題

本研究はアンケート情報を使って、推定しやすい・しにくい商品対について、また、被験者の嗜好推測の手がかりとなるようなアンケート項目について分析した。本論文で用いた手法では、画像選択の偏りと予測的中率に正の相関がみられ、偏りが大きい商品対ほど予測的中率が大きくなることがわかった。

また、アンケート情報の中では「服を購入する店」、「重視するイメージ」、「好きな

色」、「だれの目を気にするか」の項目が選好に大きな影響を与えていることが分かった。

今後の課題としては二つ挙げられる。

一つ目は、各商品対についてもう少し細かく分析する必要があるということである。選好に大きな影響を与えていた項目を見てみると、画像情報が重要であることがわかる。画像情報から商品対の違いを分析することで、よりユーザーの嗜好をとらえることができるかもしれない。

二つ目は、ビジネスへの適用を考えることである。本論文で扱った選好実験の特徴は「対比較」であること、そして「好きな方を選択する」という点である。ECサイトでは多数の商品が扱われており、その中からユーザーの嗜好に合った商品を提供する必要がある。また、ユーザーの選好を知るだけでなく、実際に商品を「購入」されることも考える必要がある。

参考文献

- 1) 殷 成久, 呉 小斌, 廣川 佐千男, 中藤 哲也, 観光イベントについての「といえば検索」の提案, 電子情報通信学会人工知能と知識処理研究会(AI), VOL. 110, NO. 301, pp.43-47 (2010).
- 2) Yamada, Y. and Hirokawa, S., Coloring For Pattern Detection, Proc. 3rd Mahasarakham International Workshop on AI, pp.27-36 (2009).
- 3) Iino, Y. and Hirokawa, S., Time Series Analysis of R&D Team Using Patent Information, Proc. 13th international conference KES 2009, Part II, Spring LNCS 5712, pp.464-471(2009)
- 4) 中藤 哲也, 大森 敬介, 廣川 佐千男, WEBDB の QUERYFORM におけるメタデータ自動抽出, DBSJ LETTERS Vol.5, No.2, pp.97-100. (2006)
- 5) リコメンデーションコンテスト2010, 主催 日本オペレーションズリサーチ学会 実践的データマイニング研究部会, <http://kgmod.jp/ormining/index.php>