

スーパーの会員アプリにおける 確率的ユーザーモデルによる推薦システムの実証実験

櫻井 瑛一^{1,a)} 本村 陽一^{1,b)}

概要: Covid19 の状況下で DX が進展したことで、デジタル空間上で人に応じたサービスを提供する基盤が整いつつある。人に応じたサービスを提供するためには、適切に人をモデル化することが重要である。我々は確率的ユーザーモデルの研究として、アンケートデータから人の類型化とその分析手段の研究を行い、ユーザーモデルを利用した情報推薦システムの作成を行ってきた。しかし、情報を推薦するだけでは最終的にユーザーの行動に影響を及ぼしたかを検証できないという弱点があった。今回、この類型に基づいた商品についてのクーポンの提示を行う推薦システムの実証実験を行い、ユーザーモデルに基づく推薦の個人への影響を検証した。

An Empirical Test of a Probabilistic User Model Recommendation System for a Supermarket App

Abstract: In order to provide person-centered services, it is important to model people. We have studied probabilistic user models as a method of modeling people, and have developed an information recommendation system using user models. However, there was a problem that we could not verify whether the user's behavior was influenced by the information recommendation alone. In this study, we conducted a demonstration experiment of a recommendation system that presents coupons based on a user model, and verified the influence of the recommendation.

1. はじめに

ものづくりからコトづくりへの転換が提言されてきているが、コトづくりを支える情報基盤の整備があまり進んでこなかった。Covid-19 の広がりによって DX 推進が叫ばれた結果デジタル化が進展し、今デジタル空間上で人に応じたサービスを提供する基盤が整いつつある。

人に応じたサービス提供を行う方法として情報推薦技術が研究され、実社会で利用されてきた。代表的な手法としては、協調フィルタリングがあり、その手法には様々なものが存在する [6-8]。協調フィルタリングは商品推薦などで使用されているが、この手法を利用するためには、構築時にはある程度の購買データが必要であること、時系列的特性(季節性商品など)への対応、データがない事象(新規商品など)への対応など様々な課題が存在する。これらを解決し、具体的な商品から多様なサービスの分野でも適用

可能にするためには、どのような傾向・特徴があれば人に響くかをシミュレーション可能なように顧客をモデル化することが重要となる。このようなモデル化はサービス分野におけるデジタルツインとして考えることができ、新規商品・サービス変更などの仮想的な介入に対する顧客の行動をデジタル空間上でシミュレーションし介入の効果検証が可能となるため、今後重要になっていくと考えられる。従って、人をモデル化する手段と作成されたモデルの検証が重要といえる。

我々は人のモデル化手段として、確率的な手法に基づいたユーザーモデルの研究を行ってきた。特に感性を反映したユーザーモデルとして、人が持つ感性に基づいた人の類型化を行い、ユーザーセグメントの特徴を抽出し、ユーザーセグメントに応じたサービスを提供する基盤的な技術の研究を行ってきた。[5]では自動車の好みに関し自動車に乗る際のわくわく感や物への好みを収集したアンケートを確率的潜在意味解析 (Probabilistic Latent Semantic Analysis [1], 以降 PLSA と略す) によりクラスタリングすることで、各人の車への考え方・感性に基づいたユーザーモデルを作成

¹ 国立研究開発法人 産業技術総合研究所 人工知能研究センター

^{a)} e.sakurai@aist.go.jp

^{b)} y.motomura@aist.go.jp

している。その結果、わくわくする車、家庭向けの広い空間を持つ車、デザインがかわいい車など各人の感性・状況に合わせたユーザーセグメントがあることを示した。また、同様の方法を用いて、食に関する分野でも適用可能であることを示した [4]。これらの研究では各分野でのユーザーのモデル化を行っているが、複数分野を横断的に見たときに、分野間で共通したモデル化部分と各個別分野特有のモデル化部分があると考えられる。[2]では、食生活と飲酒についての間で性格部分と各分野での感性部分を分離したモデル化を行っている。その結果、答えのないものを探ることが好きなセグメントでは、食・酒でこだわりがあるセグメントになる可能性が高いことなどが示された。他には、旅と農業体験についての関係性についても同様な解析を行い、この解析方法が適用可能であることを示した [3]。

これらの研究は、データからユーザーモデルを作成しその特徴を分析する形で研究が行われてきたが、実際に分析結果通りに人が行動をしているのかに課題があった。そこで、ユーザーモデルを利用した情報推薦システムを作成し、ユーザーモデルに基づく推薦が可能かの実験を行った [9,10]。この実験では、ビールに関する展示施設において、来館者に応じたビールの情報を推薦するシステムを作成した。そして、来場者と推薦に関する分析を行った結果、ユーザーモデルに従った推薦が有効であることが示された。ただ、この推薦システムでは来館者への情報提供を行っているのみであり、推薦が来館者の実際の購買に影響をあたえたかが不明であるという課題があった。

そこで、今回茨城県を地盤とする株式会社タイヨー様と協力し、タイヨー様の提供するアプリ上から画面遷移する形で [9,10] の仕組みを利用したユーザーモデルによる推薦システムの実証実験を行った。本稿では、その仕組みと実験の結果について記す。

2. ユーザーモデルを利用した推薦システム

2.1 仕組みの全体像

今回の実験ではクーポン推薦システムを作成し、株式会社タイヨー様のスマートフォンアプリ(タイヨーアプリ)を利用して実験を実施した。このアプリでは、ポイント情報管理・来店スタンプ・各店舗のチラシ情報とクーポン情報などの配信を行っている。今回はアプリのお知らせ部分で告知を行い、アプリ利用者へ実験への参加を呼びかけた。実験では、ユーザーモデルとして酒類の好みと食生活に応じたユーザーセグメントを作成した上で、セグメントごとにクーポンを推薦しその利用率の検証を行った。セグメントの作成は [5] などの分析で使用していた方法を利用して作成した。そして、推薦に用いたシステムは、[9,10]での情報推薦システム同様に POSEIDON-N1 を利用した。クーポンの提示・利用までのユーザーとシステムの情報の流れは図 1 の通りである。

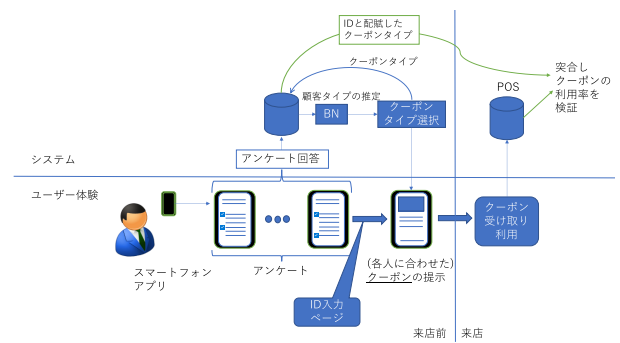


図 1 システム全体の流れ

Fig. 1 Sketch of the recommendation system

最初に、ユーザーにシステムからアンケート設問を提示し、ユーザーが回答を入力する。システムは、ユーザーセグメントの推定を行うベイジアンネットワークにアンケート回答を入力し、アンケート回答者のユーザーセグメントの推定を行う。そして、システムはセグメントに対応するクーポンの送付を行う。最後に、ユーザーはクーポンを受け取り・実際の商品を購入しそのクーポンを利用する。

2.2 ユーザーモデルの推定に用いたベイジアンネットワークの作成とアンケート設問の設計

図 1 に含まれるユーザーモデルとそれを推定するためのベイジアンネットワークとアンケートの設計について述べる。ここで用いた各手法についての紹介は次節にて行う。

ユーザーモデルの作成:

ユーザーモデルはアンケート回答が似た人をセグメント化することで作成を行った。作成に使用したアンケートは食生活と飲酒に対する価値観について問うたインターネット調査データを利用した。セグメントの作成は [5] などの分析で使用していた PLSA [1] を用いた。潜在クラス数の決定は AIC によって行った。その結果、潜在クラス数、つまりユーザーセグメント数が 4 のときが最良のモデルであることがわかった。

ベイジアンネットワークの作成:

ユーザーセグメントを推定するベイジアンネットワークをアンケート回答を元に作成した。このベイジアンネットワークから判明した 4 つのユーザーセグメントの特徴は表 1 のようになった。各々のユーザーセグメントに推薦する物は対象商品購入でポイント還元を行うクーポンとした。クーポンの対象商品はユーザーセグメントの特徴とアンケートデータからわかった各ユーザーセグメントに所属する人が日常的に飲む酒類の傾向から決定した。

アンケート設問の設計:

得られたベイジアンネットワークにてユーザーセグメントの親ノードとなったものが、セグメントの推定に強く影響を及ぼす設問である。このアンケート設問・

表1 ユーザーセグメントの特徴とクーポン対象商品

Table 1 Characteristics of user segments and recommended products for each user segment

	セグメントの特徴	クーポン対象商品
S1	外食をせず、焼酎を好む	高アルコール系飲料
S2	様々な酒を嗜み、泡系のものが好き	ハイボール
S3	家族で楽しみたい・手料理をする	プレミアムビール
S4	食事に興味があまりない、 単体で飲みやすいお酒が好み	カクテル系飲料

選択肢を含むアンケートを作成した。実験では、アンケート設問として次の9設問を設定した。

- 自宅での食事
- 食生活
- 好きな酒の傾向
- 飲酒頻度
- 普段の飲むお酒のカテゴリ
- 飲酒傾向の最近の変化
- 酒類の購買場所
- タイヨ様での購入利用理由
- 来店時チェック情報

2.3 使用した方法の紹介

2.3.1 PLSA

PLSA は文書分類で提案された手法である [1]。このモデルでは、各文書を d 、単語を w 、潜在クラスを z としたときに、文書と単語の同時確率を以下のように計算するモデルである:

$$P(d, w) = \sum_z P(z)P(d|z)P(w|z).$$

この同時確率を利用し全文書データに対する生成確率を考え、各パラメータ $P(z), P(d|z), P(w|z)$ を EM 法を用いて推定する。

今回のユーザーモデル作成では、このモデルをアンケートデータに対して適用した。この場合、文書が一人の回答者のアンケート回答に対応し、アンケートで選択された選択肢が単語に対応する。このとき、 z は回答傾向が似たセグメントとなり、このモデルからユーザーセグメントを推定することが可能となる。

2.3.2 ベイジアンネットワーク

ベイジアンネットワークは、確率変数間の条件付独立性の関係をグラフィカルモデルとして表現したものである。同時確率を条件付確率の積で分解したときに、条件側を親ノードとしてグラフを構築したものがベイジアンネットワークとなる。例えば、 X_1, \dots, X_4 の同時確率が $P(X_1)P(X_2)P(X_3|X_1, X_2)P(X_4|X_3)$ と分解されたときには、 $X_1 \rightarrow X_3, X_2 \rightarrow X_3, X_3 \rightarrow X_4$ の3つの矢印が変数間に引かれることになる。

表2 アンケートへのアクセス人数とクーポン提示までの到達率

Table 2 The number of people accessing the recommendation system and the rate at which they reach the recommendation page

	人数 (人)	割合
全アクセス数	7381	100%
最初で回答をあきらめた人	261	3.5%
途中で回答をあきらめた人	193	2.6%
ポイントカードを持っていなかった人	817	11%
推薦の提示数	6110	83%

表3 推薦されたクーポンの利用率

Table 3 Coupon usage rate for each user segment

	総数	クーポンの受け取り		クーポン利用	
		人数	割合	人数	割合
S1	152	40	26%	27	68%
S2	828	161	19%	100	62%
S3	4327	817	19%	524	64%
S4	803	159	20%	111	70%
総計	6110	1177	19%	762	65%

ユーザーモデルのユーザーセグメント推定モデルの作成と分析では、この関係性自体もデータから学習している。ベイジアンネットワークの学習には産総研が開発した Bayonet [11] を利用した。

3. 実験結果

我々は実験を 2021 年 11 月 1 日から 30 日の間行った。

アンケートへのアクセス人数、途中でのアンケート離脱人数、および推薦であるクーポン提示画面まで到達した人数の集計結果は表 2 のようになった。そして、クーポンを提示された人が、実際にアプリでクーポンを受け取り、使用した割合の各ユーザーセグメントごとの集計結果は表 3 のようになった。

各ユーザーセグメントごとに、クーポンの受け取り理由、利用理由が異なる可能性が高い。そこで、クーポンの受け取り・クーポンの利用へ影響が強い要素をアンケート回答を元にベイジアンネットワークを利用して推定を行った。その結果が表 4 から表 7 である。ここで、各アンケート選択肢の右に記入された positive/negative は、それぞれ、アンケート選択肢を選択することでクーポンの受け取り、または、クーポンの利用の確率が上昇する/減少することを意味している。

4. 考察

アプリの利用者で、アンケート回答を最後まで回答して行っていた方は全アクセス数の 90% 近くであった。これは過去のビール記念館での実験 [9, 10] と比較すると非常に高い。このような差が生じた原因は、ビール記念館では開かれた空間でサイネージ端末を使用して実験を行っ

ていたが今回は個人所有のスマートフォンを利用しているという実験環境の差、および、ビール記念館では情報提示のみであったが、今回の実験ではあらかじめインセンティブが配布されることがわかっていたという二点と考えられる。クーポン取得率は平均 20%であったが、タイヨー様によれば全アプリユーザーに普段配信を行っているクーポンの取得率と比較すると格段に高いとのことであった。このような高い取得率であったのは、自分自身で回答した結果得たという認識が影響したのではないかと予想される。

ユーザーセグメント S1 から S4 の間での取得率に差があるかを検定すると、S1 と全体の間に 5% の有意水準で差があることがわかった。このことから、S1 に属する人にはクーポンの対象とした商品が有効であったと考えられる。表 4 から表 7 の結果から、S2-S4 のセグメントでは、来店時にクーポン情報を確認するという項目が共通してクーポン取得の確率を上昇させることがわかる。クーポンを確認する習慣の割合は酒類の好みと独立と考えられるため、S2-S4 のクーポン取得率がおおよそ 20% で同じになっていた原因はクーポン対象商品に合うユーザー像であることに加え、このクーポン確認習慣ではないかと考えられる。

利用割合についてみると各セグメントの間に有意差はなかった。表 4 から表 7 の結果を見ると、S1 については、クーポン選択確率に影響する設問としてユーザーの特徴が現れ、S2-S4 の人ではクーポン利用に影響する設問としてユーザーの特徴が現れるという差があるが、どのユーザーセグメントも表 1 のユーザー特徴とほぼ一致している。従って、その特徴に良く合っていた人が商品を購入したといえる。一方、利用確率には通常時の購入場所が現れるセグメントが存在する。特に S1 では他のスーパーで酒類を買う人という特徴が、S2 ではコンビニエンスストアで他のスーパーで酒類を買う人という特徴が現れている。このことは、クーポンを出すことで普段別の場所でお酒を買う人を新たな顧客としたという行動変容を生じさせたことを意味する。

5. まとめ

インターネット調査データに基づき酒類の嗜好に関するユーザーモデルを作成し、そのモデルを利用した推薦実験を行った。その結果、ユーザーモデルの特徴に従った人々が推薦を利用することがわかり、確率的ユーザーモデルによる推薦の有効性を示すことができた。そして、推薦をだすことで、通常とは異なる場所での購入という行動変容につながる様子も確認された。

謝辞 本研究は国立研究開発法人科学技術振興機構 (JST) の研究成果展開事業「センター・オブ・イノベーション (COI) プログラム」, JPMJCE1311, の支援によって行われた。本研究における推薦プログラムの作成には、国立研究

開発法人新エネルギー・産業技術総合開発機構 (NEDO) 委託事業「人間と相互理解できる次世代人工知能技術の研究開発」の支援により作成された「POSEIDON-N1」を用いて作成された。本研究にてデータ収集に協力していただきました、株式会社タイヨー様およびサントリー酒類株式会社様この場を借りて感謝の意を表します。

参考文献

- [1] Hofmann, T.: Probabilistic latent semantic analysis, *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., pp. 289–296 (1999).
- [2] 三浦まりあ, 櫻井瑛一, 本村陽一: 展示施設来場者の確率的ユーザーモデル作成と情報推薦システムへの応用, 日本行動計量学会第 47 回大会予稿集 (2019).
- [3] 三浦まりあ, 櫻井瑛一, 本村陽一: 農業関連施設への訪問と旅行に共通した感性モデルの構築に向けて, 日本行動計量学会第 48 回大会予稿集 (2020).
- [4] 櫻井瑛一, 本村陽一, 安松 健, 道田奈々江: 食についての感性モデルに向けて, 日本行動計量学会第 46 回大会予稿集 (2018).
- [5] 櫻井瑛一, 本村陽一, 安松 健, 坂本和夫, 道田奈々江: 感性に基づく確率的ユーザーモデルの構築, 日本行動計量学会第 45 回大会予稿集 (2017).
- [6] 神脇敏弘: 推薦システムのアルゴリズム (1), 人工知能, Vol. 22, No. 6, pp. 826–837 (2007).
- [7] 神脇敏弘: 推薦システムのアルゴリズム (2), 人工知能, Vol. 23, No. 1, pp. 89–103 (2008).
- [8] 神脇敏弘: 推薦システムのアルゴリズム (3), 人工知能, Vol. 23, No. 2, pp. 248–263 (2008).
- [9] 櫻井瑛一 荒木茂樹, 安松 健, 坂本和夫, 道田本村 陽一: 展示施設来場者の確率的ユーザーモデル作成と情報推薦システムへの応用, 人工知能学会「社会における AI」研究会第 33 回研究会 (2018).
- [10] 櫻井瑛一 荒木茂樹, 安松 健, 坂本和夫, 本村陽一, 道田奈々江: ビール記念館におけるデジタルサイネージによる来場者へのリコメンデーションとその反応について, 日本行動計量学会第 47 回大会予稿集 (2020).
- [11] 本村陽一: ペイジアンネットワークソフトウェア BayoNet, 計測と制御, Vol. 42, No. 8, pp. 693–694 (2003).

表 4 S1 に分類された利用者のクーポン取得とクーポン利用に関するアンケート選択肢

Table 4 Questionnaire choices related to coupon acquisition and coupon use for users classified as S1

クーポン取得		クーポン利用	
酒の好み:アルコール度数が高め	negative	普段酒類購入場所:タイヨー各店舗	negative
普段飲む酒類:焼酎	positive	普段酒類購入場所:その他スーパー	positive
普段酒類購入場所:酒屋	positive		

表 5 S2 に分類された利用者のクーポン取得とクーポン利用に関するアンケート選択肢

Table 5 Questionnaire choices related to coupon acquisition and coupon use for users classified as S2

クーポン取得		クーポン利用	
酒の好み:強炭酸	positive	普段飲む酒類:ハイボール	positive
普段飲む酒類:チューハイ・サワー	positive	普段酒類購入場所:コンビニエンスストア	positive
普段飲む酒類:発泡酒・新ジャンル	positive	食生活価値観:ホームパーティをする	negative
食事価値観:料理は自分にとって仕事	negative		
来店時チェック情報:アプリ(クーポン)	positive		

表 6 S3 に分類された利用者のクーポン取得とクーポン利用に関するアンケート選択肢

Table 6 Questionnaire choices related to coupon acquisition and coupon use for users classified as S3

クーポン取得		クーポン利用	
酒の好み:アルコール度数が低め	negative	酒の好み:よく売れている・人気がある	positive
普段飲む酒類:プレミアムビール	positive	食事価値観:料理はできるだけ手作りに	positive
普段酒類購入場所:タイヨー各店舗	positive	食事価値観:料理は基本的に苦手だ	negative
飲酒頻度:ほとんどない	negative	食生活価値観:外食することはほとんどない	positive
来店時チェック情報:アプリ(クーポン)	positive		

表 7 S4 に分類された利用者のクーポン取得とクーポン利用に関するアンケート選択肢

Table 7 Questionnaire choices related to coupon acquisition and coupon use for users classified as S4

クーポン取得		クーポン利用	
酒の好み:アルコール度数が低め	positive	酒の好み:料理をおいしく食べられる	negative
飲酒頻度:ほとんどない	negative	酒の好み:飲みやすい	positive
当店で購入理由:ポイントが貯まるから	positive	食生活価値観:食事時にお酒を飲むことあり	negative
来店時チェック情報:アプリ(クーポン)	positive		