論文・事例研究

ファッションECサイトにおけるアンケートを用いたブランド推薦システム

田村　悠，吉住　宗朔，福永　峻，三宅　聡一郎，片山　翔太，中田　和秀

1. はじめに

近年，日本のファッションEC業界の市場規模は毎年約10%の成長を遂げ，急激な拡大を見せている [1, 2]．その裏側ではファッションECサイトが爆発的に増えたことにより，熾烈な顧客争奪競争が繰り広げられている．そこで頻繁に取られている戦略として，魅力的なクーポンの配布や，高性能な商品推薦システムの構築がある．しかしながら，現在実装されている多くの商品推薦システムでは，顧客の購買履歴が十分に蓄積された状態下でのみ顧客の嗜好を適切に捉えた推薦が可能である．そのため，購買履歴のない顧客への推薦は難しく，新規顧客の獲得やそのリピート化に対して十分な効果を発揮できない．

本研究では，顧客のアンケート回答からお勧めのブランドを推薦するシステムを提案する．この推薦システムでは，購買履歴がない顧客に対して，推薦が可能となる．その際，できるだけ精度が高く，ブランドの偏りが小さくなるような推薦を行う．偏りを小さくすることにより，ユーザーに対しては，推薦ブランドが人気ブランドばかりの退屈な推薦システムとなることを回避できる．サイト運営者に対しては，推薦されないブランドを扱っているショップのECサイト離れを防ぎ，ECサイトのさらなる規模拡大に貢献できる．推薦システムの構築には，精度向上のため，ニューラルネットワークと画像認識分野で提案された転移学習を応用した手法を利用する．また，偏りが小さい推薦のため，新たに推薦ブランド割当問題を提案し，それの最適解を用いる．そして，経営科学系研究部会連合協議会主催，平成28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数値実験により，提案システムの有効性を検証する．

たむら　ゆう，よしずみ　しゅうさく，ふくなが　たかし，みやけ　そういちろう，かたやま しょうた，なかた かずひで
東京工業大学工学院経営工学系
〒 152–8552　東京都目黒区大岡山 2–12–1
受付 17.7.25　採択 17.11.20

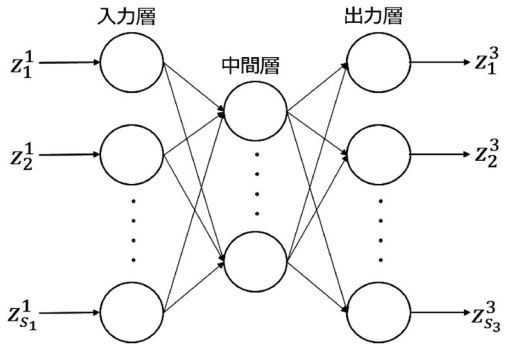

図 1　ニューラルネットワークの例

2. 既存研究

本節では，提案システムの構築に必要となるニューラルネットワークと転移学習について紹介する．

2.1 ニューラルネットワーク

ニューラルネットワーク (NN) は図 1 のように複数のユニットをもつ3種類の層，入力層，中間層，出力層で構成される．それぞれのユニットは一つ前の層に含まれるユニットから値を受け取り，一つの出力を計算する．入力層を $l = 1$，中間層を $l = 2$，出力層を $l = 3$ と表す．ここで，層 l のユニット数を s_l とし，各ユニットの出力値を縦に並べたベクトルを $\boldsymbol{z}^l \in \mathbb{R}^{s_l}$ として定義する．また，層 $l-1$ と層 l 間の結合重みを $W^l \in \mathbb{R}^{s_l \times s_{l-1}}$，層 l のバイアスを $\boldsymbol{b}^l \in \mathbb{R}^{s_l}$ とする．加えて，ベクトル値関数を $\boldsymbol{f}^l(x_1, ..., x_{s_l}) = (\sigma(x_1), ..., \sigma(x_{s_l})) \in \mathbb{R}^{s_l}$ とする．ただし，関数 σ は活性化関数と呼ばれる非線形関数である．このとき層 l の出力 $\boldsymbol{z}^l$ は次のように計算される．

$$\boldsymbol{z}^l = \boldsymbol{f}^l(W^l \boldsymbol{z}^{l-1} + \boldsymbol{b}^l)$$

出力の計算に必要なパラメータ $W^l, \boldsymbol{b}^l$ の学習には多

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

くの場合，勾配降下法を用いた誤差逆伝播法が利用される．パラメータの学習後に得られる中間層の出力値 z^2 は，ある入力に対する予測に有用な特徴をそのユニット数の次元に変換した特徴ベクトルと意味づけられる [3].

近年，画像認識の分野で中間層の数を増やし，特殊な層を利用した畳み込みニューラルネットワーク (CNN) が脚光を浴びている．しかしながらこのモデルは層の数を増やしたことで学習パラメータが多くなり，入力データが少ない場合，学習が困難になるという欠点が存在する．そのため，学習パラメータの削減が可能な転移学習の研究が行われている．

2.2 転移学習

転移学習とは新規タスクに有効な仮説を効率的に見つけ出すために，一つ以上の別のタスクで学習した知識を援用する手法である [4]．この転移学習の枠組みの一つであり CNN の精度向上を目的として提案された Fine-tuning [5, 6] について紹介する．

Fine-tuning では，データ数が十分にある予備タスクと，データ数が少なく予備タスクと入力項目が共通である目標タスクを扱う．まず，その予備タスクに対し CNN を学習する．次に，予備タスクで学習した CNN の前半部分のパラメータ（たとえば入力層と中間層の間のパラメータ）を，目標タスクの CNN に転移させる．目標タスクの学習において転移させたパラメータは学習せず固定し，残りのパラメータのみを学習する．

上記のように前半のパラメータを転用することで精度向上が図れるのは CNN のある性質が関係していると考えられている．CNN では入力層から出力層に近づくにつれて，徐々に画像の視覚的な特徴からタスク固有の意味的特徴を生成する [7]．そのため，下位層は普遍性をもち，異なるタスク間で共有することが出来る．実際に Fine-tuning を利用することで画像認識精度が向上したことが報告されている [5, 8].

3. 提案システム

顧客のアンケート回答からブランドの推薦を行う提案システムは，次の三つの要素技術で構成される．「転移学習を応用したアンケート回答からのブランド購買予測」，「回答項目を必要最低限のものにするアンケート項目削減」，「予測結果から推薦ブランドを決定するためのブランド割当最適化」である．本節では，これらの手法について順に説明する．

3.1 アンケート回答に基づいた購買予測

今回提供されたデータではアンケートは任意回答である．そのため，購買履歴が 103,144 人分存在するのに対し，アンケート回答は 3,144 人分しかない．したがって，単純にアンケート回答だけから次回購入するブランドを予測するモデルを設計した場合，アンケート回答数の不足から顧客の購買規則を正しく学習することは難しい．一方，多数の購買履歴から顧客の購買規則を学習することは容易である．そこで，多数の購買履歴から学習した正確な購買規則を転用することで，少数のアンケート回答からの高精度な購買予測を試みる．

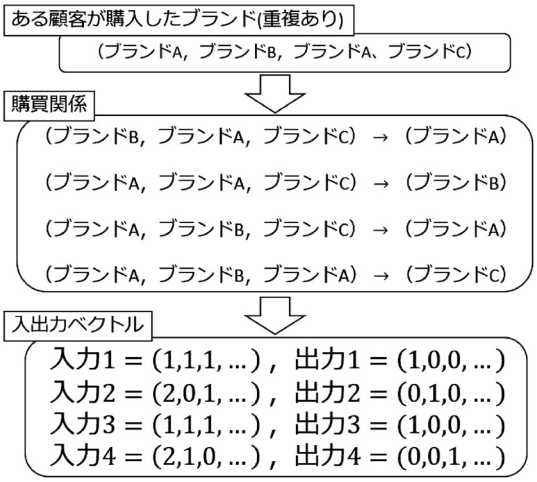

図 2 学習用データセットの作成例

提案手法は，2.2 節で紹介した Fine-tuning のアイデアに基づき設計している．予備タスクを購買履歴からの購入ブランドの予測，目標タスクをアンケート回答からの購入ブランドの予測とした．予備タスクの学習用データセットは図 2 で示した手順で作成した．図 2 では，ある顧客が購買したブランドの列（重複あり）に対し，一つのブランドを抜き出し残りのブランドと抜き出したブランドとの関係を作る．残りのブランドを入力，抜き出したブランドを出力として，各ブランドの購買数を要素とする購買ベクトルを生成する．これにより，一人の顧客から，購買数分の入出力が作られる．このデータセットによって，一人の顧客が共通して購買しているブランドの関係を，その購買数を利用して学習することができる．

本研究では，両タスクで利用する NN の中間層を Fine-tuninng と異なり，一つとしている．これは，予備タスクにおいてできるだけ正確に購買規則を学習させるためである．層の数を変化させたときの適合率[1]の変化を図 3 に示す．図 3 の実験では，4.1 節記述のデー

[1] 推薦したブランドのうち購入されていたブランド数 ÷ 合計推薦数で計算する．

 Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

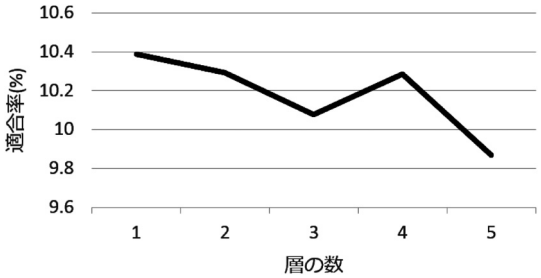

図 3　購買履歴からの推薦タスクにおける層の深さと適合率の関係

タを利用し，活性化関数を leaky-relu（関数の詳細は 4.2 節に記述），ユニット数を 15, 20, 15, 25, 20, 15 とする層を左から順に中間層に追加していった場合の適合率を計算している．図 3 より，層の数を増やすと適合率が低下することがわかる．Fine-tuninng と状況が異なる理由は明確ではないが，データ数の差やデータの特性（画像データと購買履歴）が影響を与えているものと思われる．

次にパラメータの転移方法について説明する．Fine-tuning とは逆に，提案手法では予備タスクと目標タスクの入力層が異なり出力層が同じという特徴をもっている．よって，予備タスクで学習させた後半のパラメータ（中間層と出力層の間のパラメータ）を転移させる．そして，目標タスクの学習では，転移させた後半のパラメータは固定し，前半部分のパラメータのみを学習する．パラメータ転移の流れを図 4 に示す．

目標タスクでは，特定ブランドの購入を意味する中間層次元の特徴ベクトルを，アンケート回答から予測するという学習をしていると解釈することができる．

NN のパラメータ $W^2 \in \mathbb{R}^{s_2 \times s_1}, \boldsymbol{b}^2 \in \mathbb{R}^{s_2}, W^3 \in \mathbb{R}^{s_3 \times s_2}, \boldsymbol{b}^3 \in \mathbb{R}^{s_3}$ に対し，目標タスクでは前半の二つのみを学習すればよい．今回使用したデータでは，アンケート項目は 108，ブランド数は 1064 である（すなわち $s_1 = 108, s_3 = 1064$）．つまり，転移学習によって学習するパラメータ数が 1/10 程度に抑えられる．その結果，少数のデータでも学習が可能となった．加えて次節で述べるアンケート項目削減を行うことでさらに学習パラメータ数を減らすことができる．

3.2　Ibrahim の変数重要度を用いた項目削減

アンケート項目が多い場合，回答者の回答意欲が低下し，推薦システム利用者数の減少に繋がる可能性がある．今回提供されたアンケートデータでは項目数が 108 もあり，簡単に回答できない量であった．そのため，システムを構築する際に重要度の高いアンケート項目を調べ，必要最低限の項目に絞ることにする．

表 1　提案手法の精度検証結果

手法	適合率（%）
工夫なし	6.61
転移学習	7.46
転移学習＋項目削減	7.53

NN を用いた入力変数の重要度計算に関する研究に [9] や [10] などがある．これらの手法の多くは，学習前後で重みの変化が激しいパラメータに繋がっている変数を重要度の高い変数と捉えている．しかしながら提案システムではその特性から，出力側のパラメータには学習前後で変動がない．そのため，本研究では出力側の変化量の情報を利用しない手法である Ibrahim の変数重要度 [11] を適用した．以下で Ibrahim の変数重要度の計算方法について紹介する．

入力層のユニット i と中間層のユニット j を結合するパラメータを考え，その学習後の値を W^2_{ji}，初期値を $\widetilde{W}^2_{ji}$ と表記する．このとき，i 番目の入力変数の重要度は次のように計算される．

$$RI_i = \frac{\sum_{j=1}^{s_2}(W^2_{ji} - \widetilde{W}^2_{ji})^2}{\sum_{i=1}^{s_1}\sum_{j=1}^{s_2}(W^2_{ji} - \widetilde{W}^2_{ji})^2}$$

RI_i は i 番目のユニットに接続されているパラメータの変化量が全体のパラメータの変化量に対して占める割合を意味している．この値が高い入力変数は予測に貢献している重要な変数と捉える．本研究では，パラメータの初期値として NN の標準的な実装であり python のライブラリである chainer のデフォルト値[2]を使用する．

アンケート項目の削減は入力層のユニットの削減となるため，学習パラメータ数の減少による精度向上も期待できる．転移学習や項目削減によるパラメータ削減の有効性を検証するため予備実験を行った．表 1 に転移学習や項目削減を伴って目標タスクを学習し，検証用データを推薦したときの適合率を示す．ここでは，出力層の各ユニットに格納されている各ブランドの購入度合いを意味する予測値のうち，値が最も高い五つのブランドを推薦している．

表 1 より，転移学習と項目削減の両方を行った場合の適合率が最も高く，転移学習や項目削減によるパラメータの削減は精度を向上させるうえで有効であることが確認できた．

[2] 平均 0，分散 $\frac{1}{\sqrt{s_2}}$ の正規分布からランダムに生成する．

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

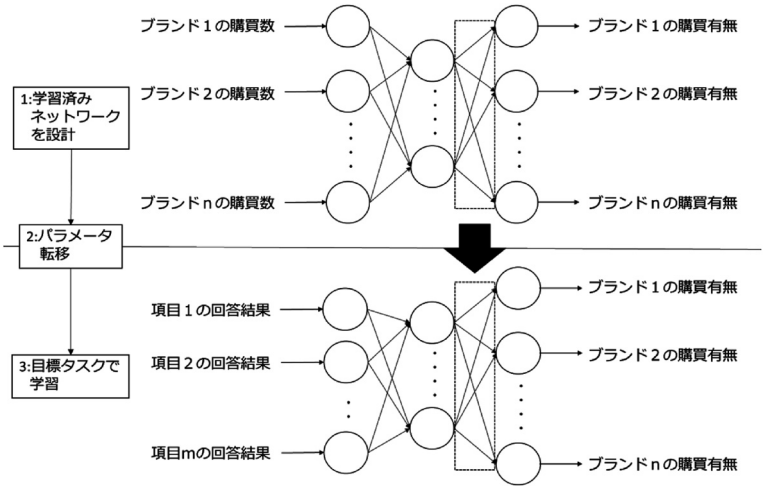

図 4　提案手法の流れ

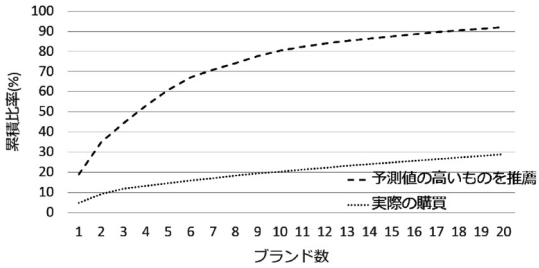

図 5　推薦数，購買数上位 20 ブランドの累積比率の比較

3.3　推薦ブランドの割当最適化

何かの商品推薦モデルに従い商品を推薦する場合，推薦する商品が売れ筋商品に偏るという傾向は頻繁に観察される．

本研究でも，前述の手法に基づき，最も購入しそうなブランドを推薦した場合，推薦するブランドが上位数十ブランドに集中してしまうという現象が見られた．その偏りを図 5 に示す．

図 5 は推薦数，購買数上位ブランドの累積比率を表している．これより，購買数上位 20 ブランドの購入数は，全体の購入数の約 30%であるにもかかわらず，推薦数上位 20 ブランドの推薦は全推薦の約 90 %も占めていることが確認できる．偏りの大きい推薦は，推薦されないブランドを埋没させ，そのブランドを扱っているショップの EC サイト離れを招く可能性がある．また，顧客側からしても意外性のないブランドを推薦されることが多くなり，精度は高いかもしれないが退屈な推薦システムとなってしまう．本研究では，これを防ぐため，各ブランドの被推薦数をコントロールしながら，購買数の期待値を最大化する推薦ブランド割当問題を提案する．なお，ここでは NN の出力は各ブランドの購買確率に相当しているものと考える．顧客 u にブランド b を推薦するとき 1，しないとき 0 を取る変数を X_{ub} としたとき，推薦ブランド割当問題として次の最適化問題を考える．

$$
\begin{aligned}
&\max \sum_{u\in U}\sum_{b\in B} P_{ub}X_{ub} \\
&\text{s.t.} \sum_{b\in B} X_{ub} = R \qquad \forall u \in U, \\
&\qquad S_b \le \sum_{u\in U} X_{ub} \le T_b \qquad \forall b \in B, \\
&\qquad X_{ub} \in \{0,1\}.
\end{aligned}
$$

ここで，U は顧客の集合，B はブランドの集合，$P_{ub} \in \mathbb{R}$ は顧客 u がブランド b を購入する確率，$R \in \mathbb{N}$ は一人の顧客に推薦するブランド数，$S_b, T_b \in \mathbb{N}$ はブランドごとの被推薦回数の下限と上限である（$0 \le S_b \le T_b \le R|U|$ とする）．

この推薦ブランド割当問題はダミーノードを導入することで図 6 のような最小費用流問題として表せる．ここで，$u_i \in U$ は特定の顧客，$b_j \in B$ は特定のブランド，D はダミーノードを表す．この最小費用流問題のすべての需要量，供給量，枝容量は非負整数であるため，すべての枝の流量が整数となる最小費用流が存在する [12]（定理 5.4）．したがって，推薦ブランド割当問題の 0-1 制約は，最小費用流問題では枝の容量の上限が 1 という条件に置き換わることに注意する．

この問題はネットワーク構造を利用することで少ない計算量で解くことができる．ここでは 実務的な状況を想定し，推薦ブランド数 R とブランド数 $|B|$ は定数として考え，顧客数 $|U|$ に対する計算量を考えることにする．

 Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

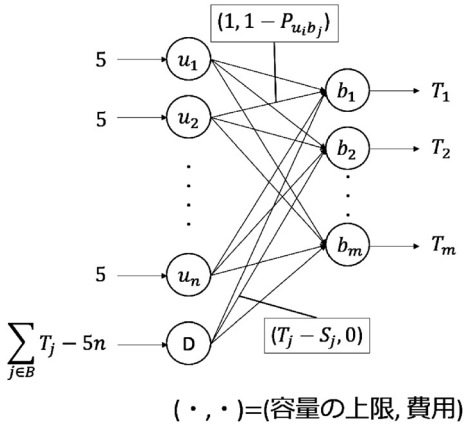

図 6　推薦ブランド割当問題のネットワーク形式

定理 1.　顧客数 $|U|$ に対し，推薦ブランド割当問題は $O(|U|^2 \log |U|)$ で解くことができる．

証明 1.　図 6 で表される最小費用流問題を逐次最短路法で解く場合，計算量は需要量（供給量）の総和 × 最短路の計算量で抑えられる [13]（9.7 節）．ここで，

$$\text{需要量の総和} \leq R|U||B|$$

が成り立つため，需要量の総和は $O(|U|)$ である．また，頂点数を $|N|$，枝数を $|E|$ としたとき，負閉路をもたないグラフの最短路は Dijkstra 法により $O(|E| + |N| \log |N|)$ で計算することができる [13]（4.9 節）．最小費用流問題のグラフは 2 部グラフであり，$|N|, |E|$ とも $O(|U|)$ である．よって，最短路は $O(|U| \log |U|)$ で計算できる．以上より，$O(|U|^2 \log |U|)$ で最小費用流問題の最適解を求めることができる．逐次最短路法では最小費用流問題の整数最適解を得るため，これが推薦ブランド割当問題の最適解に相当する．　□

定理 1 は，顧客数が増えた場合でも，推薦ブランド割当問題の求解にあまり時間がかからないことを示しており，この最適化は実用的であると考えられる．

4. 実験結果

本節では 3 種類の実験結果を述べる．まず，3 節で説明した提案システムの有用性を検証するため，ほかの購買予測手法や項目削減手法を用いた場合と比較する．次に，項目削減によりブランドの嗜好と関係の深い項目が選別できているか考察する．最後に，推薦ブランドの割当最適化の有効性を確認する．

表 2　学習データの抽出方法

	購買回数（学習期間）	顧客数
予備タスク		計 9,710 人
アンケート非回答者	15 回以上	7,962 人
アンケート回答者	1 回以上	1,748 人
目標タスク		計 1,748 人
アンケート非回答者	×	0 人
アンケート回答者	1 回以上	1,748 人

表 3　検証データの抽出方法

	購買回数（検証期間）	顧客数
推薦結果の検証		計 1,209 人
アンケート非回答者	×	0 人
アンケート回答者	1 回以上	1,209 人

4.1 実験データ

本研究で対象とした実験データは，ファッション EC サイトにおける 2015 年 4 月から 2016 年 3 月までのブランド購買履歴，その間に行われた 108 項目のアンケート回答データ，全顧客の性別，年齢データからなる．ブランド推薦システムの学習データと精度検証用データは，学習期間を 2015 年 4 月から 2016 年 1 月の 10 カ月，検証期間を残りの 2 カ月としている．

表 2 は学習データの抽出基準と抽出人数を示している．予備タスクの学習では，サイトでの購買が少ない顧客は有用な特徴が抽出できないと考え，購買回数が 15 回以上の顧客を利用する．ただし，購買回数の下限の設定については，改善できる可能性がある．アンケート回答者も同様の基準での抽出を考えたが，対象人数が激減してしまうことから，1 回以上購買のある顧客を選択した．目標タスクでは，学習の入出力に必要なアンケート回答結果と 1 回以上購買履歴をもつ顧客を抽出している．表 3 も同様に検証データの抽出基準と抽出人数を示している．検証期間に一度も購買していない顧客は推薦した結果が検証できないため除去している．なお，学習データの顧客とは重複していない．

4.2 ハイパーパラメータ

それぞれの NN の構築に必要なパラメータは次のように設計した．

- 中間層のユニット数：15
- 活性化関数：leaky relu

中間層のユニット数は学習するパラメータ数を減らすため，小さい値で設計した．活性化関数には，近年頻繁に利用される leaky relu を適用した．leaky relu は

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

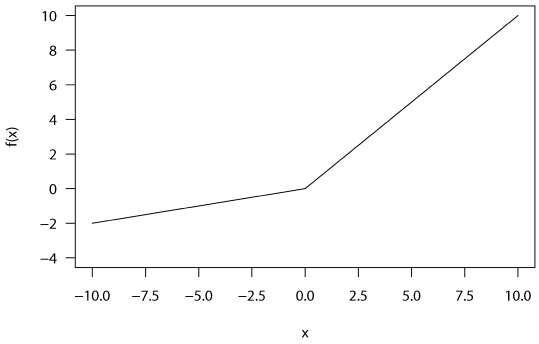

図 7　leaky-relu 関数のグラフ

以下の式で表され，図 7 のような関数である．

$$f(x) = \max\{0.2x, x\}$$

leaky relu による出力結果は $[0, 1]$ に入らない．そのため，以下の式にて確率化している．

$$\tilde{z}_i^3 = \frac{z_i^3 - \min_k z_k^3}{\sum_k^{s_3} z_k^3}$$

なお，活性化関数の選択として，ほかの関数の利用も考えられるがそれは今後の課題である．変数重要度で絞る項目数は顧客の回答意欲を削がない値として 15 項目にした．

比較対象とした L2 正則化付きロジスティック回帰モデル[3]とランダムフォレストの実装には python の機械学習ライブラリである scikit-learn を使用し，ハイパーパラメータにはデフォルト値を用いた．ただし，ランダムフォレストでは精度向上のため，デフォルト値から木の数，葉ノードに含まれるデータ数を 100, 5 と変更した．ほかに比較手法として用いた転移学習を利用していない NN のハイパーパラメータは，前述の NN のハイパーパラメータと同じものを利用した．各比較手法の学習には 3.1 節の提案手法にある予備タスクの学習は行わず，目標タスクの学習のみ行った．

各ブランドの被推薦数はおおむね実際の購買数に比例させることを目標とし，各ブランドの被推薦数が占める割合を実際の購買数の割合と等しくなるようにした被推薦数に対し，その 0.85 倍と 1.25 倍を下限 S_b と上限 T_b にしている．なお，下限，上限に利用した値は実際の適用場面に応じて，適宜変更可能である．

4.3　実験結果

まず，3 節で提案した推薦システムの有効性を検証するため，ほかの購買予測手法や項目削減手法との比較実験を行った．ここでは，購買予測手法として L2 正則化付きロジスティック回帰モデル，ランダムフォレスト，NN，転移学習を用いた NN を使用する．また，項目削減手法として，ランダムフォレスト学習の際に計算されるジニ係数[4]と提案手法である Ibrahim の変数重要度を用いる．これらの数値実験の結果を表 4 に示す．この数値実験では，項目削減手法を適用した後，削減後の項目を入力変数として購買予測を行う．その結果に対して，推薦ブランドの割当最適化を行い，その適合率を比較している．適合率以外に頻繁に利用される評価指標に再現率[5]，正解率[6]があるが，これらの指標は手法間の精度比較をした際，大小関係が適合率と一致する．そのため，本研究では適合率の検証結果のみで比較している． 全体を通して，転移学習を用いた NN，L2 正則化付きロジスティック回帰モデル，ランダムフォレスト，NN の順に適合率が高かった．これは，今回の状況のような少ないデータ数に対して，転移学習を用いた NN による推薦の有効性を示唆している．なお，最も適合率の高かった Ibrahim の変数重要度で項目削減をし，転移学習を用いた NN を使った場合，推薦ブランド数が 5 のとき，再現率：7.24%，正解率：99.26%，推薦ブランド数が 10 のとき，再現率：13.1%，正解率：98.82%であった．次に，アンケート項目削減手法の比較を行う．提案システムでは，ジニ係数を利用するよりも Ibrahim の変数重要度を用いて項目削減をするほうが適合率が高く，提案システムの

表 4　各手法の精度（適合率）検証結果

手法	変数削減	推薦ブランド数（%）	
		5	10
ロジスティック回帰	ジニ係数	4.34	3.81
	Ibrahim	4.52	4.03
ランダムフォレスト	ジニ係数	4.09	3.77
	Ibrahim	4.07	3.72
NN	ジニ係数	3.87	3.64
	Ibrahim	4.00	3.73
転移学習を用いた NN	ジニ係数	4.38	4.07
	Ibrahim	**4.56**	**4.13**

[3] 推定するパラメータを θ，$L(\theta)$ をロジスティック回帰の誤差関数，正則化パラメータを C としたとき $L(\theta) + \frac{1}{2C}\sum_j \theta_j^2$ を最小化することで学習を行う．本研究では $C = 1$ とした．

[4] ジニ係数は各クラスの分散の総和として定義される．変数重要度はランダムフォレストのアルゴリズムにより複数の決定木を学習させた際に，各説明変数の分割によるジニ係数の減少度をすべての木で平均することで算出される．

[5] 推薦したブランドのうち購入されていたブランド数 ÷ 実際の購買ブランド数で計算する．

[6] （推薦したブランドのうち購入されていたブランド数 + 推薦しなかったブランドのうち購入のなかったブランド数）÷（顧客数 × ブランド数）で計算する．

 Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

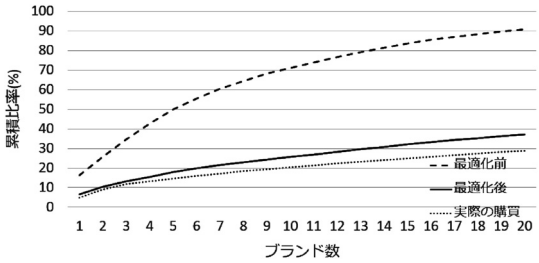

図 8　最適化後の推薦数上位 20 ブランドの累積比率

表 5　削減後に残ったアンケート項目（重要度が高い順）

項目カテゴリ	項目内容
参加イベント	音楽フェスに参加したか
ファッション観	芸能人や有名人の服装を真似るか
ファッション観	新しい商品に敏感か
ファッション観	原産国や素材を気にするか
ファッション観	伝統のブランドに精通しているか
ファッション観	特殊なファッション観があるか
ファッション課題	流行が分からない
意識	衝動買いをするか
ファッション観	服は個性を発揮するための手段か
ファッション観	何かファッション観があるか
ファッション観	従来のスタイルにこだわるか
ファッション観	ファッションは個性表現か
人生重視価値観	競争・勝利
ファッション課題	ファッション課題があるか
ファッション観	特定の好きなブランドがある

有効性が確認できる．L2 正則化付きロジスティック回帰モデルでも Ibrahim の変数重要度を用いたほうが優れていることから，推薦手法によらず本質的に重要な項目が選択できていると考えられる．ランダムフォレストではジニ係数を利用したほうが若干適合率が高いが，これはジニ係数の計算にランダムフォレストの結果を用いているため，相性がよかったものだと考えられる．

次に，Ibrahim の変数重要度によりブランドの嗜好と関係の深い項目が抽出できているか確認する．表 5 は選択された項目を重要度が高い順に並べている．アンケート 108 項目のうちファッション観カテゴリに属すものは 32 項目である．選択された 15 項目のうち，ファッション観カテゴリに属するものは 10 項目と多数を占める．ファッション観に属する項目はブランドの好みと関係が深いと考察でき，重要な項目が多く抽出できていることがわかる．最後に，推薦ブランドの割当最適化の有効性検証のため，推薦ブランド数を 5 としたときの最適化後の推薦ブランドの累積比率とブランドの推薦数，適合率の変化について確認する．推薦

表 6　最適化による推薦数，適合率の変化

	各ブランドの被推薦数の順位			
	1〜20 位	21〜40 位	41〜60 位	61 位〜
推薦数（個）				
最適化前	5,026	665	106	248
最適化後	2,247	909	634	2,255
適合率（%）				
最適化前	8.59	3.01	3.77	1.61
最適化後	8.90	4.18	2.52	1.06

数上位 20 ブランドの累積比率を図 8 に示す．図 8 より，最適化後の累積比率は最適化前と比較して実際の購買に近づき，目的に沿った推薦が実現できていることが確認できる．各ブランドの被推薦数の順位が 1〜20 位，21〜40 位，41〜60 位，61 位〜のブランドの各推薦数と適合率を表 6 に示す．表 6 より，最適化することで推薦数 1〜20 位，21〜40 位の人気ブランドにおいて過剰な推薦が抑えられ適合率も向上していることが確認できる．一方で推薦数 41〜60 位，61 位〜のブランドにおいて推薦数が適切になったものの適合率が低下している．これは，精度を上げることのみを目的とすると人気ブランドを多く推薦してしまい，推薦数を適切にするには最適化が必要であることを示している．

5.　おわりに

本稿ではアンケート回答を用いたブランド推薦システムを提案した．システム構築にあたって，アンケートデータからの購買予測，アンケート項目削減，推薦ブランドの割当を行った．それぞれの段階において，購買予測では Fine-tuning を応用した手法を提案，項目削減では提案手法の特徴を考慮し，Ibrahim の変数重要度を適用した．推薦ブランドの割当ではブランド数の偏りが少ない推薦を行う最適化問題を提案し，その問題を最小費用流問題に帰着することで高速に求解可能であることを示した．

アンケートから推薦を行うシステムの実用例に，日本電気株式会社による AI 活用味覚予測サービスというものがある [14]．このサービスは，簡単なアンケートに回答すると，好みのうまい棒の味が予測，推薦される．これは先進的かつ手軽に利用できるという理由から一時話題となった．提案した推薦システムでも，インターネット上で誰もが利用可能とするなどにより同様の効果が期待できる．その結果，利用者数とともに会員数が増え，新規顧客の獲得に貢献できると考えられる．

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

また，本研究ではブランド推薦割当問題を用いて実際の購買割合に沿った推薦を試みている．上限と下限を状況に合わせて設定することにより，顧客全体の購買期待値を上げつつ，特定の商品の販売促進を行うような推薦も可能である．このように，提案したブランド推薦割当問題の応用性は広いと思われる．

謝辞　松井知己氏から有益なコメントをいただきました．ここに記して感謝いたします．データを提供していただきました，データ解析コンペティション関係者の皆様にお礼を申し上げます．

参考文献

[1] 経済産業省,「平成 27 年度我が国経済社会の情報化・サービス化に係る基盤整備（電子商取引に関する市場調査）」, http://www.meti.go.jp/press/2016/06/20160614001/20160614001-2.pdf （2017 年 9 月 19 日閲覧）
[2] 経済産業省,「平成 28 年度我が国におけるデータ駆動型社会に係る基盤整備（電子商取引に関する市場調査）」, http://www.meti.go.jp/press/2017/04/20170424001/20170424001-2.pdf （2017 年 9 月 19 日閲覧）
[3] 神嶌敏弘,『深層学習』, 近代科学社, 2015.
[4] 神嶌敏弘,“転移学習,” 人工知能学会誌, **25**, pp. 572–580, 2010.
[5] R. Girshick, J. Donahue, T. Darrell and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, 2014.
[6] G. Wimmer, A. Vécsei and A. Uhl, “CNN transfer learning for the automated diagnosis of celiac disease, ” In *Proceedings of 2016 Sixth International Conference on Image Processing Theory, Tools and Applications*, 2016.
[7] 中山英樹,“深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習,” 信学技報, **115**(146), pp. 55–59, 2015.
[8] P. Agrawal, R. Girshick and J. Malik, “Analyzing the performance of multilayer neural networks for object recognition,” In *Proceedings of European Conference on Computer Vision 2014*, pp. 329–344, 2014.
[9] G. D. Garson, “Interpreting neural network connection weights,” *AI Expert*, **6**, pp. 47–51, 1991.
[10] J. D. Olden and D. A. Jackson, “Illuminating the ‘black box’: A randomization approach for understanding variable contributions in artificial neural networks,” *Ecological Modeling*, **154**, pp. 135–150, 2002.
[11] O. M. Ibrahim, “ A comparison of methods for assessing the relative importance of input variables in artificial neural networks,” *Journal of Applied Sciences Research*, **9**, pp. 5692–5700, 2013.
[12] R. K. Ahuja, T. L. Magnanti and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, 1993.
[13] 繁野麻衣子,『ネットワーク最適化とアルゴリズム』, 朝倉書店, 2010.
[14] NEC,「AI 活用味覚予測サービス by NEC the WISE」, http://jpn.nec.com/bigdata/aiprofiler/ （2017 年 7 月 21 日閲覧）

 Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.