

# EC サイトにおけるソーシャルログイン会員のサイト内行動分析 —ゴルフ用品 EC サイトを対象とした会員の特徴と購買行動分析—

大竹 恒平<sup>†1</sup> 橋本 鴻<sup>†2</sup> 生田目 崇<sup>†1</sup>

**概要**：本稿では、ゴルフ商品を扱う EC (Electronic Commerce) サイトを対象に、ソーシャル会員と一般登録会員それぞれが行うサイト内行動の特徴とその差異について評価することを目的とする。具体的には、会員登録データ及びアクセスログデータから作成した説明変数を用い、決定木分析を用いた判別モデルを作成した。会員変数を用いた決定木分析の結果、ソーシャルログインサービスへの登録は、サイト内でのロイヤルティに基づき行われるわけではなく、EC サイトへのロイヤルティが低い、新規会員が多数を占めることが分かった。また、行動変数を用いた決定木分析の結果、購買に至るセッション時の滞在時間において、ソーシャル会員と一般会員のサイト内行動の差異の一端が示された。

**キーワード**：消費者行動、ソーシャルログイン、決定木

## 1. はじめに

近年、ストレージ容量の拡大やデータ処理技術の革新によるビッグデータ時代を背景に、様々な分野で大量かつ多様なデータが取得されるようになった。こうしたデータ活用が最も進んでいる分野がビジネス領域である。特に、幅広いメーカー・商品カテゴリの取り扱いがあり、多様な顧客を有する、EC (Electronic Commerce) サイトにおいては、ビッグデータを活用したマーケティング活動が求められている。

このような状況の下、学術領域においては、消費者行動やチャネル最適化などマーケティング領域に関する研究が盛んに行われている。中でも、消費者の購買行動に着目した研究は、産学問わずに高い関心を集めている。例えば、サイト上の購買履歴とアクセスログを用い、購買生起確率を推定した研究 [5] や、サイトの閲覧履歴から、特徴的な購買パターン発見並びに顧客クラスを作成し、重要顧客の発見を試みた研究 [4]、ドラッグストアの ID 付き POS (Point of Sales) データを用いて特徴的な購買パターンを特定した研究 [1]、EC サイトへの訪問者が購買前に取った閲覧行動から、購買のタイミングを予測するモデルを構築した研究 [2]、消費者の複数カテゴリに跨る購買行動に注目し、価格プロモーションと購買との関係のモデル化並びに、ドラッグストアの ID 付き POS データを用いてモデルの実証分析を行った研究 [9] などがある。

他方で、近年、インターネット上における情報発信・共有ツールとして、ソーシャルメディアが広く利用されている。ソーシャルメディアとは、「ブログ、ソーシャルネットワークワーキングサービス (以下 SNS)、マイクロブログ、動画共有サイトなど、利用者が情報を発信し、形成していくメデ

リア」を指し<sup>1</sup>、その利用者数は年々増加しており、生活者誰もが情報の受発信ができることが特徴として挙げられる。これらの普及の背景には、その利用の手軽さに加え、スマートフォンの普及に伴う携帯端末の多機能化、モバイルアプリケーションの登場などによるユーザビリティの向上や、コミュニケーションの変容などが考えられる。

上記に示す一連の流れを背景に、ソーシャルメディアを利用した広告・マーケティング活動が注目されている。ソーシャルメディアは、商品の紹介から活用方法、レビューなど、企業の製品・サービスの広告メディアとして活用されている一方で、商品開発・販売戦略の策定や、自社の顧客特性の解明に必要な、ライフログに基づく詳細な行動データの収集場所としての活用も期待されている。

また、ビジネス領域でのソーシャルデータの活用を目的とした取組みの一つとして、ソーシャルログインがある。ソーシャルログインとは、アカウント連携機能 (シングルサインオン) の一種であり、ソーシャルメディア上の ID を用いて、企業の EC サイトやポータルサイトにログインする機能を指す。ソーシャルログインを用いることで、EC サイト上の顧客 ID と SNS アカウントの連携ができ、従来は知りえなかった、自社の顧客が有する興味・関心や商品評価などの情報が消費者単位で取得可能となる。

筆者らはこれまで、上記のソーシャルログインへの登録を行った会員 (以下、ソーシャル会員) に着目し、ソーシャル会員の購買行動に関する研究に取り組んできた [6]。具体的には、Facebook 及び Twitter のソーシャル会員及び一般会員が 1 年間に行った購買履歴を ID 付き POS データから抽出し、商品カテゴリをノード、商品間の併売関係をエッジとするネットワークグラフを作成し、ネットワーク構造の比較を行った。その結果、ソーシャル会員は一般会

<sup>†1</sup> 中央大学 (連絡先: otake@indsys.chuo-u.ac.jp)

<sup>†2</sup> 中央大学大学院

投稿: 2016 年 11 月 30 日

採択: 2017 年 2 月 10 日

1 総務省 2013 年 6 月報道資料

員に比べ、併売されやすい商品同士の組み合わせが存在し、これらの商品は、季節性を含んだ商品カテゴリ（サマーアパレル、ウインターアパレル）並びに、高価格商品カテゴリ（ドライバー、フェアウェイウッド等のゴルフクラブ類）であることが明らかになった。これまでの筆者らの研究では、主に ID 付き POS データから得られた、消費者の購買履歴（結果）を用いて分析を行った。他方で、ウェブの閲覧履歴（アクセスログデータ）や会員情報（デモグラフィックデータ）については、活用をしていない。今後、EC サイトにおいては、ソーシャル会員の増加が見込まれており、上記のデータを活用したサイト内行動の特徴を明らかにする必要がある。

ソーシャルログインは、前述したようにアカウント連携サービスであり、SNS を定期的に利用している、即ち SNS のヘビーユーザの会員が登録していると考えられる。総務省の調査によれば、ソーシャルメディアの利用者の目的として、「情報探索」が上位に挙げられ、全年代で見た場合、約 4 割の利用者が目的として挙げている [8]。以上のことから、ソーシャル会員は、キュレーションサービスを含んだ、インターネット上での情報探索能力の高い会員であると考えられる。この様な視点に立つと、ソーシャル会員はウェブ上での比較・検討、流入経路、閲覧回数といった、EC サイト上で購買に至るまでのサイト内行動が一般会員と異なることが想定される。

## 2. 研究目的

本稿では、EC サイトにおいてソーシャル会員と一般登録会員それぞれが行うサイト内行動の特徴とその差異について評価することを目的とする。具体的には、アクセスログデータ及び会員登録データを用い、会員属性やサイト内行動を表現することを期待した説明変数を作成し、決定木分析を用いた判別モデルを作成する。前述の様に、ソーシャル会員が有する情報探索能力は、一般会員に比べて高いことが想定される。そこで、判別モデルの重要度及び係数の値から、ソーシャル会員と一般会員の探索行動の比較を行い、ソーシャル会員のサイト内行動への特徴付けを行う。なお、本稿ではゴルフ商品を扱う EC サイトを対象とするが、先行研究の結果を参考に、事前にソーシャル会員をクラブ、アパレル、用品・小物の 3 カテゴリに分類し、それぞれのカテゴリに対して判別モデルを作成する。これにより、特性が異なる商品カテゴリごとに、ソーシャル会員の探索行動に対する特徴付けを行うことが出来る。なお、モデルの作成・検証に当たり、株式会社ゴルフダイジェスト・オンライン社2（以下、GDO）にデータ提供などの協力をい

ただいた。

本稿の構成について記す。第 3 節では、分析に利用するデータの概要について示す。第 4 節では、本稿の分析の流れを述べる。第 5 節では、データセット及び判別モデルに利用する説明変数を示す。第 6 節では、決定木分析により作成した、ソーシャル会員と一般会員の判別モデルについて、評価及び考察を述べる。第 7 節では、ソーシャル会員と一般会員それぞれにおいて、購買に至るまでの探索行動の判別モデルについて、評価及び考察を述べる。最後に、本稿の結論及び今後の課題についてまとめる。

## 3. データ概要

本稿で利用するデータは、ID 付き POS データ、会員登録データ、アクセスログデータの 3 種類である。ID 付き POS データには、商品名、商品金額、購買日など、商品に関するデータが含まれている。会員登録データには、職業、性別、年齢、ゴルフスコアなどの会員の属性が含まれている。アクセスログデータには、GDO の運営するサイトへのアクセス回数、利用デバイス、ページ滞在時間や回遊数などが含まれている。なお、運営サイトではゴルフ関連商品の EC サイトに加え、ゴルフ場予約サイト、ニュースサイト、スコア管理などの会員サービスなどが含まれる。

本稿では 2014 年 7 月 1 日～2015 年 6 月 30 日の 1 年分のデータを対象とし、会員のデータに限定して用いた。データの概要を表 1 に示す。ただし、これらは GDO の全トランザクション及びトラフィックデータから会員のみをサンプリングしたデータを集計したものである。

表 1 本稿が対象とするデータの概要

データ期間：2014/07/01-2015/06/30		
ID 付き POS データ	購買件数(合計)	約 1,000,000
	購買者数(ユニーク)	約 210,000
アクセスログデータ	アクセス数(合計)	約 1,320,000,000
	セッション数(合計)	約 144,000,000
	アクセス者数	約 5,160,000
会員登録データ	登録者数(2015/6/30 時点)	約 2,700,000

## 4. 分析の流れ

本研究では、はじめに、全サイト会員の中から、次節で説明するようにソーシャル会員を抽出する。次に、抽出したソーシャル会員の ID 付き POS データに記録された購買商品を用い、該当期間に 1 回以上購買があった会員を分類する。さらに、それぞれのソーシャル会員の、アクセスロ

2 ゴルフダイジェスト・オンライン <http://www.golfdigest.co.jp/> (2017年2月11日筆者確認)

グデータ、会員登録データを用いて、判別モデルに利用する説明変数の作成を行う。そして、これらの説明変数を用いて、決定木分析による、ソーシャル会員と一般会員の判別を目的としたモデルの作成を行う。最後に、ソーシャル会員、一般会員の両者において、決定木分析による購買生起の判別を目的としたモデルの作成を行う。

## 5. データセットの作成及び判別モデルに利用する説明変数の作成

はじめに、データセットの作成について述べる。GDOが有するECサイトにおいて、2015年6月30日時点でFacebookアカウントを用いてソーシャルログイン登録を行っている、55,533人のソーシャル会員に関して、購買商品の抽出を行った。その結果、データ期間において、7,771人（総購買件数：35,742件）のソーシャル会員の購買があった。次に、これらの商品をキーとして、GDOが定める中分類名称（クラブ、アパレル (Men, Lady)、用品・小物）毎に会員を分類し、データセットを作成した。その結果、クラブでは2,675人（総購買件数：5,984件）、アパレルでは4,171人（総購買件数：14,624件）、用品・小物では5,113人（総購買件数：15,134件）のソーシャル会員が購買を行っていることが分かった。次に、それぞれのデータセットに分類されたソーシャル会員について、アクセスログデータ及び会員登録データの抽出を行った。以後、これらのデータセットをソーシャルデータセットと呼ぶ。ソーシャルデータセットに関するアクセスログデータの概要を表2に示す。なお、セッション (Session) 数は、ある会員のECサイトへの訪問から離脱までを「1」として集計した値を、ページビュー (Page View) 数は、1セッション中に訪問したページ数を集計した値を指す。また、コンバージョン (Conversion) 数は実際に商品の購買に至った回数を集計した値を指す。

表2 ソーシャルデータセットに関する  
アクセスログデータの概要

	クラブ	アパレル	用品・小物
セッション数	629,797	875,872	1,043,805
ページビュー数 (PV数)	8,557,246	11,658,965	13,390,577
コンバージョン数 (CV数)	5,984	14,624	15,134

次に、各商品カテゴリのソーシャル会員数と同数の一般会員を選択するようにランダムサンプリングを行い、上記に示した一連のデータセットの作成手順によりデータセットの作成を行った（表3）。以後、これらのデータセットを一般データセットと呼ぶ。

表3 一般データセットに関する  
アクセスログデータの概要

	クラブ	アパレル	用品・小物
セッション回数	842,809	1,307,052	1,470,319
ページビュー数 (PV数)	13,821,650	22,173,540	22,445,094
コンバージョン数 (CV数)	13,235	41,075	31,336

次に、それぞれのデータセットに対して、アクセスログデータ及び会員登録データを用い、サイト内行動を表現することを期待した説明変数の作成を行った。アクセスログデータからは、「PV数」「滞在時間」といった説明変数をセッション毎に集計を行った。会員登録データからは、「年齢」「サイト登録年数」といった説明変数を作成し、会員毎に集計を行った。以下の分析では、アクセスログデータから作成した説明変数を行動変数と、会員登録データから作成した説明変数を会員属性変数と呼ぶ。

## 6. 決定木分析を用いたソーシャル会員の判別モデルの作成

本節では、アクセスログデータ及び会員登録データから作成した変数の内、ユーザ毎に集計した説明変数を用い、ソーシャル会員と一般会員の判別を目的としたモデルの作成を試みる。判別モデルの作成方法としては、決定木分析を用いる。

### 6.1 決定木分析

決定木分析とは、観察対象データの集団を、従属変数に対し最も効率よく分類できる独立変数(原因)によって次々と分割し、木の枝のように分岐・整理していく分析手法である。データの集団を効率よく分類・整理できるので、ルール抽出・生成や予測モデル構築などに利用される。決定木においては、出現する独立変数をノードと呼ぶ。また、最上位にあるノードをルート、決定木中の部分木は、枝ないしブランチと呼ぶ。

決定木分析では、分岐させる属性の決定方法により、生成された決定木の性能に大きな影響を与える。昨今では、様々な決定木作成アルゴリズムが提案されているが、本稿では、カイ二乗値により分類基準を設定するCHAID (Chi-squared Automatic Interaction Detection) [3]を用いる。CHAIDでは、分岐によって分割される子ノードにおいて、目的変数であるカテゴリの構成比率に統計的な有意性があるかどうかを、カイ二乗分布で評価される分割表の検定に帰着する。分割候補の変数それぞれについて、分割を行った場合のカイ二乗値を求め、自由度から特定されるカイ二乗分布

と比較して p 値を求める。p 値の大きさを比較することで、最善の分割ルールを特定する。なお、CHAID では、CART (Classification and Regression Tree) のような 2 分木ではなく、分割される子ノードが 3 つ以上になる多分木である。

## 6.2 ソーシャル会員の判別モデルの作成

6.1 節で述べた決定木分析手法を用い、ソーシャル会員と一般会員の判別モデルの作成を行う。なお、本分析には会員単位で集計した会員属性変数を用いる。表 4 に、判別モデルの作成に用いた変数の一覧を示す。

表 4 ソーシャル会員の判別モデルに用いた説明変数の一覧

	説明変数名	データ形式
基準変数	会員種別 (ソーシャル会員 : 1, 一般会員 : 0)	2 値
会員属性変数	age (年齢)	連続値
	sex (男性 : 1, 女性 : 0)	2 値
	site_reg_count (サイト登録年数)	連続値
	dm_flag (ダイレクトメール登録 : 1, 未登録 : 0)	2 値
	score_reg_count (スコア管理登録 : 1, 未登録 : 0)	2 値

## 6.3 ソーシャル会員の判別モデルの評価及び考察

6.2 節にて述べた説明変数を用い、決定木分析によるソーシャル会員の判別モデルを作成した。なお、判別モデルの作成は、全データを対象にランダムに学習用データ : 検証用データ=3:1 と分割して学習及び精度の検証を行った。また、モデルの作成には SPSS Modeler の CHAID を利用した。CHAID アルゴリズムによる分岐選択のパラメータとしては、カイ二乗値の有意確率を 0.01 とし、親ノードは全体の 5% 以上のデータ数になる様に制約を設けた。なお、カイ二乗値の計算には尤度比を用い、Bonferroni の調整を行ったものを分岐時の基準として用いた。ソーシャル会員を 1、一般会員を 0 とし、全データを用いて作成した決定木を図 1 に示す。作成した決定木を用いた検証用データの分類を行い、予測結果の評価を行った。評価項目としては、表 5 に示すコンフィュージョンマトリックスを作成し、正解率 (式 1)、適合率 (式 2)、再現率 (式 3) を算出した。なお、正解率は、予測結果全体がどの程度事実と一致するかを示す指標である。また、適合率は、正と予測したうち実際に事実が正であった割合を、再現率は、事実が正のうち予測が正であった割合を示す指標である。一般的に、適合率と再現率はトレードオフの関係となる。

表 5 予測結果の評価に用いたコンフィュージョンマトリックス

	事実が正	事実が負
予測が正	True Positive (TP)	False Positive (FP)
予測が負	False Negative (FN)	True Negative (TN)

$$\text{正解率} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$\text{適合率} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{再現率} = \frac{TP}{TP+FN} \quad (3)$$

表 6 にソーシャル会員の判別モデルの分類精度の概要を示す。

表 6 ソーシャル会員の判別モデルの分類精度

対象データ \ 評価指標	正解率	適合率	再現率
	全データ	83%	35%

図 1 より、第一分岐に採用された説明変数は、GDO の有する EC サイトへの登録年数であった。ソーシャル会員においては、登録後 3 年以内の会員が全体の約 7 割を占めている。一方で、一般会員においては、登録後 3 年以上の会員が全体の 8 割以上を占めている。第二分岐においては、ソーシャル会員が多く含まれるノード 10、11 では年齢が、また、一般会員が多く含まれるノード 12、13 ではダイレクトメールフラグが分岐基準として採用された。年齢に着目すると、ノード 10 ではソーシャル会員の半数以上が 43 歳以下であり、以後年齢が増加するごとに割合が減っていることが分かる。ダイレクトメールフラグに着目すると、ノード 12、13 合わせて、約 9 割の一般会員がダイレクトメールなし、ゴルフスコア管理 (第 3 分岐) に登録していることが分かる。

GDO では 2013 年 7 月にソーシャルログインサービスの導入が開始されている。ただし、ソーシャルログイン登録は、既に GDO アカウントを有している会員も可能である。上記のサイト登録年数の分岐結果から、ソーシャル会員はソーシャルログインサービスの導入以後に新規に登録を行った会員が大半を占めていることが伺える。また、ソーシャル会員の分岐条件として、ダイレクトメールやスコア管理など、GDO が発信するコンテンツへの接触は確認されなかった。以上の結果から、ソーシャルログインサービスへの登録は、サイト内でのロイヤルティにより行われているわけではないこと、ソーシャル会員は、GDO サイトへのロイヤルティが高くない、新規会員であることが推察される。

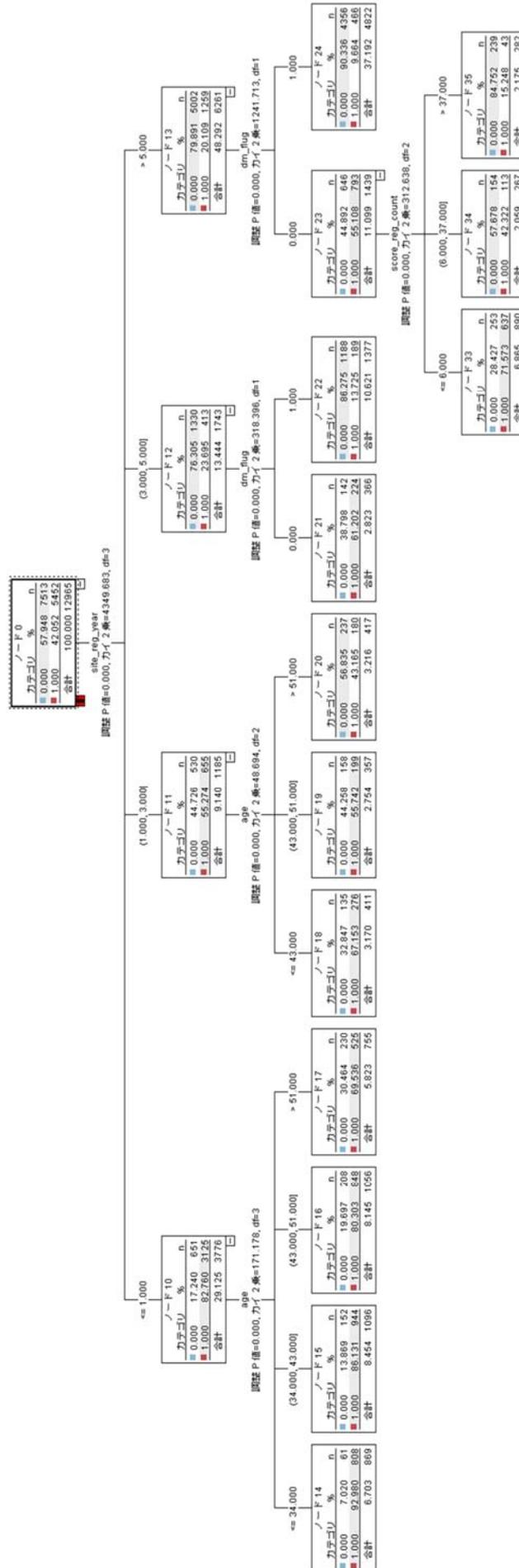


図1 全データを用いたソーシャル会員の判別モデルの結果

## 7. ソーシャル会員及び一般会員の購買生起の判別モデル

本節では、購買時の購買行動の特徴を明らかにするため、ソーシャル会員、一般会員それぞれにおいて、購買生起の判別モデルを作成する。

### 7.1 購買生起の判別モデルの作成

購買生起の判別モデルでは、それぞれのセッションにおける購買の有無を基準変数とする。また、商品カテゴリ毎に判別モデルを作成する。なお、本分析では、購買の有無を基準変数とするため、セッション単位で集計した説明変数を用いる。また、会員属性変数はいずれも、行動変数を説明変数として用いる。ただし、セッション単位で集計

を行った場合には、正解ラベル付きのセッションに対して、不正解ラベル付きのセッションの数が大幅に発生する。そこで、データセットの作成時には、正解ラベルに合わせてサンプリング処理を行った。以下に、判別モデルの作成に用いた変数の一覧を表7に示す。なお、リファラー (Referrer) は、ECサイトにアクセスする直前に閲覧していたページを指す。表7における外部リファラーサイトは、google や Yahoo! といった検索エンジンなど、GDO 外のサイトからの流入を示すフラグである。これに対し、表7における内部リファラーサイトは、ゴルフ場予約やゴルフレッスン、ニュースなどの GDO が有するコンテンツからの流入を示すフラグである。また、閲覧時間帯の3変数はどの時間帯に閲覧が行われたかを示すフラグであり、値はセッション開始時の時間に基づき定めた。

表7 購買生起の判別モデルに用いた説明変数の一覧

	説明変数名	データ形式	
基準変数	セッションにおける購買の有無 (購買 1,非購買 0)	2 値	
行動変数	外部リファラーサイト	google	2 値
		yahoo	2 値
		bing	2 値
		facebook	2 値
		rakuten	2 値
		Ameblo	2 値
		Kakaku.com	2 値
		amazon	2 値
	内部リファラーサイト	www (トップページ)	2 値
		shop (ショップページ)	2 値
		reserve (ゴルフ場予約ページ)	2 値
		lesson (レッスンページ)	2 値
		news (ニュースページ)	2 値
		usr (ユーザページ)	2 値
		myp (マイページログイン)	2 値
		style (ゴルフスタイルページ)	2 値
		sns (コミュニティページ)	2 値
		faq (ヘルプページ)	2 値
		morning:閲覧時間帯(朝:5:00~10:59)	2 値
		daytime:閲覧時間帯(昼:11:00~16:59)	2 値
night:閲覧時間帯(夜:17:00~22:59)	2 値		
pv_count (ページビュー数)	連続値		
stay_time (滞在時間)	連続値		

表 8 ソーシャル会員及び一般会員におけるそれぞれの判別モデルの精度の概要

会員種別 評価項目 対象	ソーシャル会員			一般会員		
	正解率	適合率	再現率	正解率	適合率	再現率
全データ	83%	77%	93%	81%	75%	92%
クラブ	83%	46%	92%	81%	45%	91%
アパレル	82%	45%	90%	81%	42%	84%
用品・小物	84%	45%	81%	82%	46%	92%

## 7.2 購買生起の判別モデルの評価及び考察

7.1 節にて示した行動変数を用い、決定木分析によるソーシャル会員の判別モデルを作成した。なお、判別モデルの作成は、前節と同様全データのうち 3/4 を学習用データ、1/4 を検証用データとしてランダムサンプリングを行い学習及び精度の検証を行った。前節同様 CHAID を用い、分析の設定も前節と同様である。ソーシャル会員、一般会員それぞれにおいて、購買が行われたセッションを 1、行われなかったセッションを 0 とし、全データを用いて作成した決定木を図 2, 3 に示す。また、6.3 節同様に、それぞれのモデルに関する分類精度の算出結果を表 8 に示す。図 2, 3 より、ソーシャル会員、一般会員ともに、第一分岐においては、PV 数が採用された。PV 数が多いほど、購買に至るセッション数が多いことが分かる。その実数に着目すると、ソーシャル会員では、PV 数が 12 回以上、一般会員では、PV 数が 14 回以上になると購買に至るセッション数の割合が 5 割を上回り、回数が増える毎に割合が増加する傾向にある。また、第二分岐においては、ソーシャル会員と一般会員共に、滞在時間及び内部リファラー（ショップページ）が採用された。まず、一般会員の滞在時間に着目すると、ノード 266（図 3）を親ノードとするノード 276, 277, 278 では、滞在時間が長くなるほど、購買に至るセッションの割合が高くなっていることが分かる。一方で、図 2 ソーシャル会員のノード 4 を親ノードとするノード 14, 15, 16 に着目すると、滞在時間が長くなるほど、購買に至るセッションの割合が低くなっていることが分かる。また、図 2 のノード 2 を親ノードとするノード 9, 10, 11 においても同様に、およそ 15 分（907 秒）を境に、滞在時間と購買に至るセッションの割合は負の相関関係がみられる。次に、内部リファラー（ショップページ）に着目すると、ソーシャル会員では 1 つ、一般会員では 2 つのノードが採用された。そのうち、購買に至るセッションの割合を見てみると、どちらも他のショップから経由して商品ページに訪れたセッションの方が、購買割合が高いことが分かる。

上記の結果は、ソーシャル会員と一般会員の購買行動の違いの一端を示すものであると考える。筆者らがこれまでに行った、同様の EC サイトを対象とした研究より、商品

ページにおける滞在時間が長いほど、購買の生起確率は向上することが確認されている [7]。しかしながら、ソーシャル会員においては、購買に至るセッション数の割合及び実数共に、滞在時間が短いセッション程、購買に至りやすいことを示している。前述した滞在時間と購買生起確率の関係は、消費者のサイト内での購買行動に起因すると考えられる。サイト内での購買を考えた際、複数の商品間の比較・検討や、商品の詳細情報の閲覧などの購買行動により滞在時間は増加する。一方で、滞在時間が短く購買に至る消費者は、外部サイトで何らかの探索行動を経て購買に至ることが推察される。EC サイト上での購買は、EC サイト内での比較・検討に基づき購買が行われる場合もあれば、単にネットショッピングのプラットフォームとして購買が行われる場合もある。後者の消費者であれば、比較・検討は行わず、直接商品購買を行うため、滞在時間は短くなると考える。

次に、商品カテゴリ（クラブ、アパレル、用品・小物）毎に作成した判別モデルについて考察を行う。これは、対象となる商品カテゴリによって、消費者の行動は異なると考えたためである。その結果、クラブ、アパレル、用品・小物の 3 商品カテゴリにおいても、全データを用いた場合と同じ傾向で、第一分岐で PV 数が採用されることが確認された。第 2 分岐以降に採用された滞在時間については、特にクラブ類、用品・小物類において、同様の傾向が得られた。一般的に、高価な買い物を行う際の検討時間は、安価な買い物を行う際の検討時間に比べて長くなる。前述した過去の研究結果では、ドライバー、アイアン、パターなどの高価格・耐久消費商品においては、滞在時間が長く、ティーやボールなどの低価格・消耗品においては、滞在時間が短くなる傾向が確認されている [7]。一般会員の 3 商品カテゴリにおいては、滞在時間が分岐基準に採用された場合は、滞在時間が長くなるほど、購買に至るセッションの割合が高くなる傾向にあった。一方で、ソーシャル会員のクラブ、用品・小物の 2 つの商品カテゴリにおいては、滞在時間が短くなるほど、購買に至るセッションの割合が高くなる傾向にあった。ただし、アパレルに関しては、滞在時間と購買に至るセッションの割合に大きな差は発見されなかった。この差は、商品特性によるものであると考える。

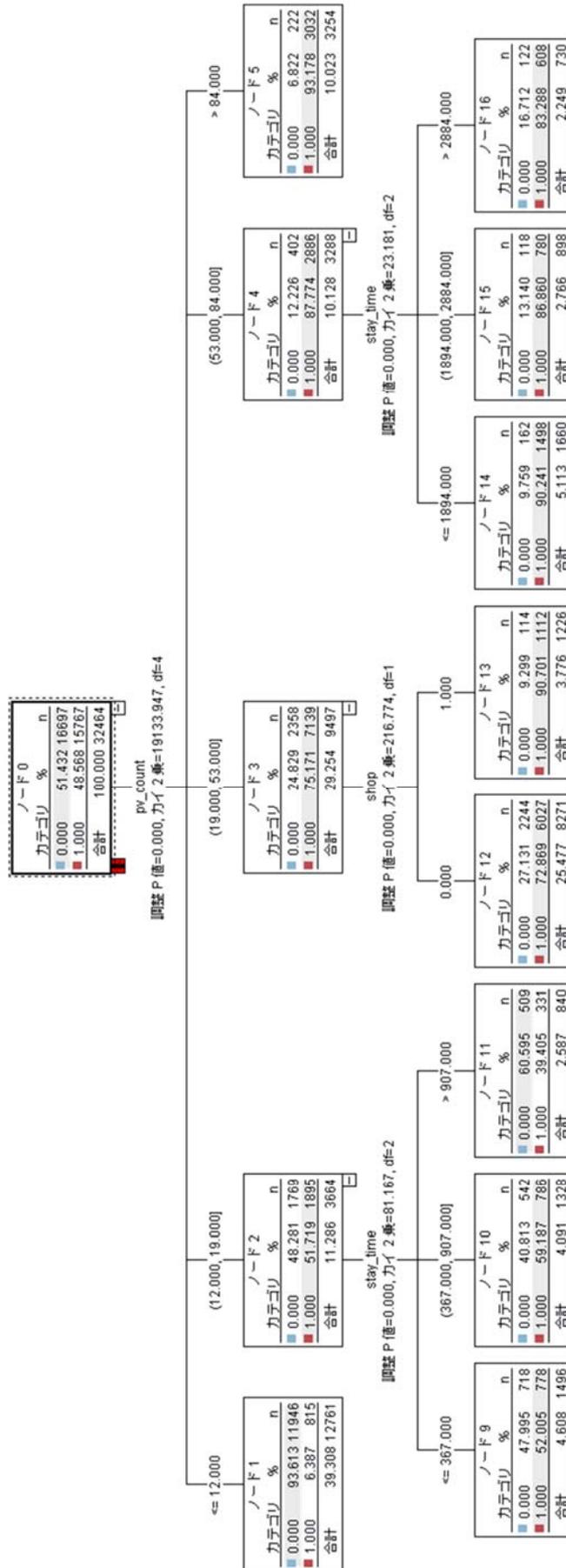


図 1 全データを用いたソーシャル会員の購買判別モデルの結果



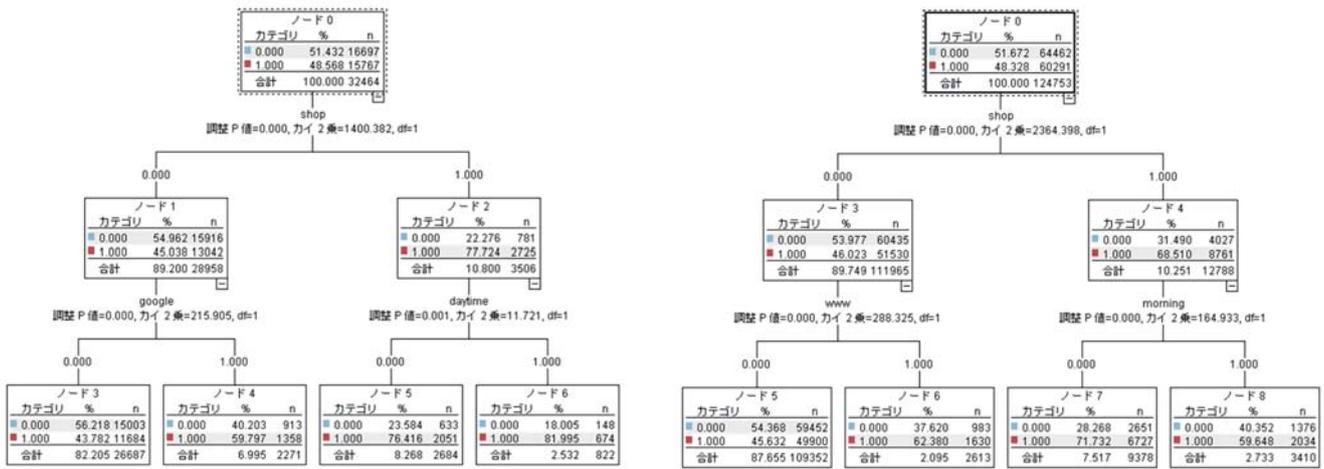


図3 PV数、滞在時間を除いた全データによる決定木の結果 (左：ソーシャル会員，右：一般会員)

まず、用品・小物に関しては、基本的に安価であり価格差も少ない。比較項目も、他の商品カテゴリに比べて少ないため、一般会員においても、滞在時間は短い。これに対し、クラブは高価格帯に分類され、価格差も大きい。加えて、商品詳細情報が細かく記載されているため、検討項目も自ずと多くなる。ただし、商品によって規格は定められており、事前にメーカーのWebサイトや他のECサイトを通じて探索行動を行っていた場合には、容易に検索から購入まで進めることが出来ると考える。一方で、アパレルは好みが反映されやすい商品であり、品揃えも多様である。生地や販売メーカー等の詳細情報以外にも、色見や会員自身の有するアイテムとのコーディネート等、個人の感性により検討する視点が異なる。そのため、滞在時間は人によって異なり、こういった個人差をモデルに反映することは困難であったと考える。

最後に、上記以外の説明変数の影響を明らかにするため、PV数・滞在時間を説明変数から除き、再度判別モデルの作成を行った。結果を図4に示す。

図4から、大枠の構造は変わらないものの、ノード1(図4左)の分類基準としてはgoogle(検索エンジン)が採用されていることが分かる。一方で、ノード3(図4右)の分類基準としては、www(GDOトップページ)が採用されていることが分かる。どちらもリファラーを示す説明変数であるが、ソーシャル会員はGoogleの検索エンジンを用いて検索を行い、GDOのECサイトへアクセスした場合の購入に至る割合が多いのに対し、一般会員では、GDOのトップページから検索を行った場合の購入に至る割合が多いことを示している。

上記のアクセスログデータから作成した説明変数を用いた、一連の決定木分析の結果から、ソーシャル会員と一般会員の購入に至るセッションにおける購買行動の差異の一端が示された。

## 8. 結びに

本稿では、アカウント連携機能(シングルサインオン)の一種であり、ソーシャルメディア上のIDを用いて、企業のECサイトやポータルサイトにログインする機能の総称である、ソーシャルログインサービスへの登録会員(ソーシャル会員)を対象に、ECサイトにおいてソーシャル会員と一般登録会員のサイト内行動の特徴とその差異について評価することを目的とした。具体的には、アクセスログデータ及び会員登録データから、会員属性及びサイト内行動を表現することを期待した説明変数を作成し、決定木分析により評価を行った。結果を以下にまとめる。

- ✓ 会員属性変数(表4)を用いた決定木分析を行った結果、ソーシャルログインサービスへの登録は、サイト内でのロイヤルティに基づき行われるわけではないこと、ソーシャル会員は、GDOサイトへのロイヤルティが高くない、新規会員が多数を占めることが推察された。
- ✓ 全商品カテゴリを対象に、行動変数(表7)を用いた決定木分析を行った結果、滞在時間に注目すると、一般会員の場合は滞在時間が長くなるほど、購入に至るセッションの割合が高くなるのに対し、ソーシャル会員の場合は滞在時間が長くなるほど、購入に至るセッションの割合が低くなっていることが分かった。これらの結果から、ソーシャル会員は外部サイトでの比較・検討を十分に行った上で当該ECサイトに訪問し、購入のみ行うといった探索行動の一端が推察された。
- ✓ クラブ、アパレル、用品・小物の3つの商品カテゴリを対象に、行動変数(表7)を用いた決定木分析を行った結果、特にクラブ類、用品・小物類において、全商品カテゴリを対象とした場合と同様の探索行動の一端が推察された。また、PV数、滞在時間を除いた説明変数で、同様に決定木分析を行った結果、外部からの流入リファ

ラーとして、ソーシャル会員では google が分岐基準に採用された。

以上の一連のアクセスログデータから作成した説明変数を用いた決定木分析の結果から、ソーシャル会員と一般会員の購買に至るセッションにおける購買行動の差異の一端が示された。

ソーシャルログインサービスの普及は急速に広まっており、今後ソーシャル会員は増加していくと考えられる。本稿により得られた知見は、今後、ソーシャル会員を対象とした企業のマーケティングアプローチの際の指針となると考える。本分析の結果から、現状ではソーシャルログインサービス自体は、あくまでユーザの利便性を向上させるものであり、ロイヤルティを表現する様な効果は確認されなかった。また、ソーシャル会員の行動変数を用いた決定木分析の結果から、ソーシャル会員は、購買時に商品情報の事前収集を行った上でECサイトに訪れるといった“賢い”消費者である可能性が示唆された。以上の結果を踏まえると、導入開始以後は、ソーシャル会員の定着化に向けた顧客ロイヤルティの向上施策を行う必要があると考える。ソーシャルログインサービスの性質上、登録時の負担は軽減されるが、反対にサイトスイッチングの障壁も下がる。ソーシャルメディア上の会員についての情報を用いて顧客の特性を理解した上で、最適な商品推薦、セール情報の配信等の広告宣伝が求められる。一部 EC サイトでは既に試されているように、ソーシャル会員に限定したクーポン配信や、値引きセール等は、ソーシャル会員を獲得する際のきっかけとしては有効であると考えられる。しかしながら、本稿の分析結果から、ソーシャル会員は、比較・検討をサイト外で行っている可能性があるため、特価商品を目玉商品とした、バンドル販売等の販売戦略には注意が必要であると考えられる。現状においては、サイト運営企業はソーシャル連携に関して十分なマーケティング活動が行われていないといえ、今後新たなチャネルとして活用することにより、ソーシャルログインのメリットを増加させることができると考えられる。

今後は、ソーシャル会員に関するソーシャルメディア上から取得可能なデータと合わせた、会員特性の分析を行う予定である。具体的には、会員が行った投稿データや、お気に入り、会員の訪れたロケーション等のデータを抽出し、ソーシャルメディア上の行動データを用いた説明変数を作成した上で、商品の購買行動との関係性について調査を行う予定である。

**謝辞** 本研究において株式会社ゴルフダイジェスト・オンライン社様からは大変貴重なデータの提供並びにディスカッションを通じた数多くの有益なコメントをいただきました。ここに謝意を表します。

## 参考文献

- [1] 久松俊道, 朝日弓未, 山口俊和, “ドラッグストアの ID 付き POS データを用いた日用品購買パターンの比較分析,” オペレーションズ・リサーチ, Vol. 57, No. 2, pp. 63-69 (2012).
- [2] 久松俊道, 外川隆司, 朝日弓未, 生田目崇, “EC サイトにおける購買予兆発見モデルの提案,” オペレーションズ・リサーチ, Vol. 58, No. 2, pp. 93-100 (2013).
- [3] Kass, G. V., “An Exploratory Technique for Investigating Large Quantities of Categorical Data,” *Journal of the Royal Statistical Society, Series C*, Vol. 29, No. 2, pp. 119-127 (1980).
- [4] 小池雄平, 菅谷健人, 住田潮, 高橋一樹, 平野智章, 山本浩平, “アクセスログデータに基づく重要顧客のリアルタイム判別,” *Department of Social Systems and Management Discussion Paper Series*, No. 117, University of Tsukuba (2007).
- [5] Moe, W. W. and Fader, P. S., “Dynamic Conversion Behavior at E-Commerce Sites,” *Management Science*, Vol. 50, No. 3, pp. 326-335 (2004).
- [6] Otake, K. and Namatame, T., “Analysis of Purchasing Tendency using ID-POS Data of Social Login User,” *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 9, pp. 116-123 (2016).
- [7] Otake, K. and Namatame, T., “Analysis of Impacts on the Purchasing Behavior Using Information on Social Networking Services,” *Proceedings of the 17<sup>th</sup> Asia Pacific Industrial Engineering and Management Systems Conference*, 6 pages (2016).
- [8] 総務省, 「情報通信白書平成 27 年度版」 (2016).
- [9] 鶴見裕之, 澁谷浩太郎, 村瀬明宏, “小売業のカテゴリー間プロモーション・マネジメント : 消費者の複数カテゴリー購買行動モデル,” オペレーションズ・リサーチ, Vol. 50, No. 2, pp. 92-98 (2005).

## **Analysis of Behavior of Social Login Customers on EC Site Through the analysis of customer characteristic and purchasing behavior about social login customers on the golf EC site**

Kohei OTAKE<sup>†1</sup> Ko HASHIMOTO<sup>†2</sup> Takashi NAMATAME<sup>†1</sup>

**Abstract:** In this paper, we aim to evaluate the characteristics and differences of the behavior within the site conducted by social login customers and general customers for EC (Electronic Commerce) site. Specifically, we provided discrimination models using decision tree analysis focused on the explanatory variables created from customer information data and access history data. From the result of decision tree analysis using customer information data, it became clear that social login customers are new customer and do not have high royalties within the site. Moreover, from the result of decision tree analysis using access history data, it became clear that part of the difference in social login customers' and general customers' behavior within the site in terms of staying time at the session leading to purchasing.

**Keywords:** Customer Behavior, Social Login, Decision Tree Analysis

---

<sup>†1</sup> Chuo University (Correspondence Author: [otake@indsys.chuo-u.ac.jp](mailto:otake@indsys.chuo-u.ac.jp))

<sup>†2</sup> Graduate School of Chuo University

Submitted: 30, Nov., 2016

Accepted: 10, Feb., 2017