



レディース・ファッションに関する インターネット上のクチコミについてのデータ解析

峰 滝 和 典

要旨 本稿では、レディースファッションに関するインターネット上のクチコミについて、主としてテキストマイニングと機械学習を用いた分析を行った。主たる結論としては、第一に抽象的で曖昧な「おしゃれ」という語句と類似度が高かったもののなかに、各年共通しているものとして「似合う」があったが、年によって結果は大きく異なった。第二に評価ポイントに関してSVM (Support Vector Machine)、決定木、ランダムフォレスト、k近傍法といった教師ありの機械学習を行った結果、5段階評価を2段階に絞った場合、70%超の正答率となった。以上の教師無しの教師ありの機械学習の各手法で用いた説明変数について、機械学習の非階層クラスター分析を行った結果、「黒」、「デザイン」、「リピーター」、「かわいい」といった語句がクラスターの形成に大きく寄与していることがわかった。

Abstract This research conducted mainly through text mining and machine learning for WOM related to ladies' fashion on the internet. The main results from the data analysis were as followings: (1) The word "fashionable," which was abstract and ambiguous, was similar to "suit" every year. However, the similar words were largely different by year. (2) The reputation points were analyzed by supervisor learning, such as SVM, Decision Tree, Random forests, and K-nearest neighbors. As a result, the accuracy rate was over 70% in the case of two categories condensed from 5 categories. The un supervisor clustering method was applied for explanatory variables used in supervisor learning, and the words "black," design," repeater," and "kawaii" were found, which contributed mainly to form clusters.

キーワード クチコミ, テキストマイニング, 機械学習, SVM (Support Vector Machine), 決定木, ランダムフォレスト, k近傍法

原稿受理日 2022年9月21日

1. はじめに

近年、インターネット上での購買活動は様々な品目に及んでいる。そのなかでファッションを取り上げたのは、人々のクチコミの影響を受けやすいと考えたことと、届いた商品がサイズや生地、色合いなどが購入者の意図通りではない可能性も考えられるのではないかと推察したからである。また季節や年によってクチコミの内容も変わってくるかどうかを検証することも本稿の目的の一つである。

クチコミに関してこれまでに膨大な研究がなされてきた。その全てを網羅することは本稿の範囲を超えている。古くは、Arndt (1967) はクチコミを、「話し手と受け手の間の対人コミュニケーションであること、ブランド・商品・サービス・店に関する話題であること、受け手が非商業的な目的であることと知覚していること」の3条件を満たしているものと定義している。濱岡 (1994) は、Arndt (1967) の3条件に加えて、「話し手と受け手が社会的な関係に規定されている」という条件を加えて、クチコミを定義している⁽¹⁾。また、杉本 (1997) は、「商品やサービスの購入に関して、消費者同士で行われる人的コミュニケーション」と定義している。

本稿が対象とするのは、「楽天データセット」のなかからレディースファッションに関するアイテムを抽出したデータである⁽²⁾。クチコミの内容についても主として商品やそれに付随する郵送などのサービスに関するものであり、他のユーザーのコメントを受けたクチコミもあるが、基本的にはユーザー自身の意見や感想を述べているものである⁽³⁾。

本稿では分析方法としてテキストマイニングや機会学習を用いるため、クチコミ自体をデータ解析した先行研究を参考にした。対象としては観光・旅行 (高野 (2009), 辻井他 (2012), 石橋 (2012), 野守他 (2015), 久保田 (2016), 松本 (2016), 大川 (2019), 合川 (2021)), 化粧品 (白田他 (2012) 高島他 (2016), 松波他 (2016), 酒井他 (2019)), 料理のクチコミ (中村他 (2018), 立花他 (2021)) に関する文献が多くみられた。そのなかから分析方法の観点からいくつか紹介する。

ホテル・旅館が提供するサービスの品質は経験財である (石橋 (2012))。旅行者は実際に宿泊するまではサービスの質は不確かであり、旅行者が過去利用者のクチコミや評価ポ

(1) この条件を加えたのは、「噂」との区別を明確にするためとある (濱岡 (1994))。

(2) 他のアイテムについても随時分析していく。

(3) 書かれたクチコミを別のユーザーが読むことや参考にするケースもあり、その意味ではコミュニケーションと本稿では解釈する。

イントをチェックする動機がここにある。石橋（2012）は形態素解析を行い、性別にどのような形容詞と名詞の出現回数が多いかをまとめている。辻井他（2012）は、エリアごとにクチコミの特徴表現と数値評価の関係を示している。クチコミに書き込まれた特徴表現を、数値評価項目ごとに肯定極性と否定極性に分類して抽出することで、宿泊検討者の利便性の向上につながることを示唆している。野守他（2015）は、PLSA（Probabilistic Latent Analysis）を適用して、観光テーマを抽出した上で、各観光地の特徴を分析している。PLSAは文書とそこに出現する単語の間には潜在的な意味クラスがあることを想定して、各文書における単語の出現頻度が記録された「文書」と「単語」の共起行列データを学習し、文書と単語の共通のトピックスとなるような特徴を見つけ出す方法である（野守他（2015））。久保田（2016）は頻出語句に対応分析を行い、単語間・ホテルグループ間の対応関係を調べて、クラスター分析（デンドログラム）を行い、ホテルをグルーピングしている。大川（2019）はWordnetの同義語集合を利用して、日本語、英語、中国語のクチコミを解析している。語のクラスターにおける類似度を分析し、観光施設に対して言語ごとの印象を分析している。施設によっては言語間で同じ印象をもたれているが、異なる印象となっているものもある（大川（2019））。

白田他（2012）は化粧品のクチコミについて、形態素解析により、名詞、形容詞、動詞を抽出したのちRIDF（Residual Inverse Document Frequency）により重みづけを計算し、特徴語として抽出している。高島他（2016）は、化粧品のクチコミについて形態素解析を行い、名詞の数、動詞の数、形容詞の数、副詞の数、体言止めの文の数を計算している。それ以外に、クチコミの文字数、改行の数、星の数、購入品/モニター、現品/サンプル、リピートの有無、画像の数、感嘆符の数、極性、話題の転換数、話題の詳細度、LDA（Latent Dirichlet Allocation）を提案素性として取り出している。極性については、ポジティブと解釈された文の数とネガティブと解釈された文の数を引いて定義している。ポジティブ、ネガティブの判断のために、教師あり学習を用いた分類モデルのひとつであるSVM（Support Vector Machine）を用いている。話題の転換数と話題の詳細度についてはクチコミの各文をベクトルに変換し、隣接する文のコサイン類似度をもとに計算している。LDAについては、トピックス数100に設定した、化粧品に特化したLDAのトピックスモデルを構築し、各クチコミが、得られたトピックスを潜在的に有する確率を計算している。最終的にクチコミの有用性判定システムとして、クチコミデータベースから素性抽出ベクトル化し、ベクトル化されたクチコミと正解ラベルとなるLike数を回帰モデルで学習させ、決定係数とRMSE（Root Mean Square Error）により評価している。松波

他(2016)はコスメアイテムに関するレビューを分析し、コスメアイテムに対する評価項目別の辞書を構築し、それに基づいた自動スコアリング実験を行っている。酒井他(2019)は、化粧品のレビューから評価表現を抽出し、レビューの評価値をもとに評価表現に対する評価値を算出し、評価項目ごとに分類している。

立花他(2021)は、明確に定義できていない味覚表現の一つであるコクに注目し、料理レシピサイトにおけるクチコミ分析を行っている。まずコク表現とサッパリ表現を二値としたテキスト分類を行い、TF-IDF(Term Frequency Inverse Document Frequency)を特徴量とするSVM(Support Vector Machine)を用いて、機械学習による分類を行っている。また機械学習モデルの結果を解釈する手法であるLIMEを用いて、コク表現に寄与する特徴語として食材に関する用語(ごま油、オイスターソース、チーズ、バター、マヨネーズ、黒糖、砂糖、卵黄)が多く確認されている。

以上でみたように、テキストマイニングはまず形態素解析を行う。それ以降は、比較的容易な方法である、頻出語句の集計や共起ネットワークや階層的クラスターを描いたりする記述的分析から機械学習を用いたSVM(Support Vector Machine)やLDA(Latent Dirichlet Allocation)等のトピックス分析、LIMEまで、様々な手法が試みられてきた。

本稿では2において、月別の頻出語句の集計、3においてWord2Vecを用いたコサイン類似度を用いた分析、4では統計分析のひとつである順序プロビットモデルによる評価点分析、5では教師あり機械学習を用いた評価点に関する分析の評価、6で機械学習(教師なし)を用いた非階層クラスター分析、7で本稿での分析からわかったことをまとめた上で今後の課題を述べたい。

2. 抽出語句の月ごとの比較

先述したように本稿では、「楽天データセット」のなかからレディースファッションに関するアイテムを抽出したデータをもとに分析を行う。用いる変数は評価点、クチコミ期間としては2015年1月から2019年12月までの月別データである。レディースファッションは、季節要因、年ごとの流行が考えられるので、基本的には月別の分析を行った。次別に形態素解析を行った結果を紹介する⁽⁴⁾。形態素解析には、Python上、MeCabを利用した。

(4) ブランクのあるものや、csvファイルに変換したときに列単位に区切られていないものを除いたクチコミを対象とした。その結果分析の対象となる年毎の総クチコミ数は、2015年1,834,492、2016年784,525、2017年639,105、2018年623,552、2019年607,777となった。

コンピューター解析によって、文章を最小単位の品詞に分解し、単語の品詞の種類、活用形の種類などを割り出していき、単語ごとに全文書数における出現度、全文書数とその単語が含まれる文書数の割合、単語の分散表現などを可能とする。

本稿では形態素解析を行った結果のうち、月別に分析したい単語をいくつかのカテゴリーに分類して頻出回数をクチコミ数合計で除して100をかけた値を図表1に掲載した。カテゴリーは、「色」、「美」、「華やか・豪華」、「かわいらしさ」、「視覚に関する用語」、「季節感」、「場・機会」に関する語句群に分けた⁽⁵⁾⁽⁶⁾。図表1を見てカテゴリーごとに月別にみた特徴をまとめる。「色」は6月を挟んだ2カ月、つまり4月から8月に値が大きい傾向にある。「美」は3月、4月に多くみられる。「華やか・豪華」は値自体が他のカテゴリーとして少ないが、1月から3月に相対的に多い傾向にある。「かわいらしさ」は7月、8月と多くなっている。「視覚に関する用語」は、7月前後が多い傾向にある。「季節感」は年によってばらつきがあるが、傾向として3月4月と夏に多くみられる。「場・機会」は3月4月と、7月から9月と年2回のピークがある。レディースファッションに関するクチコミのなかに、こうした季節的な特徴があることがわかった。形態素解析によって頻出語句を用いた分析は記述的で決して分析レベルは高くないが、ファクトファインディングを積み上げることで、実務面で役立つこともあると考える。今後さらに分析を深めたい。

(5) 各カテゴリーに入っている語句は以下の通り。

	色	美	華やか・豪華	可愛らしさ	視覚に関する用語	季節感	場・機会
アイボリー	ブラウン	キレイ	エレガント	かわいい	一目	サマー	イベント
イエロー	ブラック	きれいな	ゴージャス	かわいらしい	映える	夏	オフィス
エメラルド	ブルー	美しい	シック	可愛い	丸見え	夏向き	お呼ばれ
オレンジ	ベージュ	美容	セクシー	可愛らしい	鏡	夏場	お出かけ
カーキ	ホワイト	綺麗	艶		見え	夏物	お正月
カラー	マスタード		華やか		見える	寒い	お盆
カラフル	ミント		華やかさ		見かける	初夏	クリスマス
クリーム	モスグリーン		豪華		見せる	寒い	クリスマス
グレー	ライム				見た目	暑い	ジム
レッド	ラベンダー				後ろ姿	真夏	スポーツ
ゴールド	レッド				撮る	雪	公園
サーモン	レモン				撮影	残暑	行事
シルバー	ローズ				写真	秋	祭り
ダーク	黄色				写す	秋口	七五三
チャコール	黒				透ける	冬場	謝恩
パール	黒い				露出	冬物	書月
パステル	黒色					肌寒い	通学
ビーチ	黒地					猛暑	通勤
ピンク	紺						同窓会
ピンク色	紺色						二次会
水色	真っ黒						披露宴
	真っ白						保育園
							法事
							北海道
							幼稚園

(6) 他にも、「アイテム」、「用途別分類」、「素材・生地・織り方・模様・デザイン等」、「品質」、「流行」、「コーディネート」、「清楚」、「カッコいい」、「着心地・触感」、「感謝・喜びの気持ち」、「楽しい気持ち」、「アイテムに関するネガティブな語句」、「その他ネガティブな語句」、「身体」、「大きさ・スタイル」、「金銭」、「衣服以外のアイテム」というカテゴリーを作成した。

図表1 語句カテゴリーの時系列推移

2015年	色	美	華やか・豪華	可愛らしさ	視覚に関する用語	季節感	場・機会
1月	42.8	8.6	1.7	27.7	18.9	14.8	9.0
2月	44.3	9.8	2.0	27.8	20.2	16.8	11.3
3月	44.3	10.4	2.0	28.2	21.0	18.6	11.8
4月	44.8	10.2	1.6	27.8	21.7	19.3	11.6
5月	44.3	9.7	1.4	27.9	22.4	19.4	10.0
6月	45.4	9.3	1.3	29.5	23.0	18.6	11.0
7月	43.2	8.5	1.2	33.5	22.3	16.8	11.9
8月	41.5	8.6	1.3	33.8	22.1	16.6	11.6
9月	41.8	9.4	1.5	27.6	20.8	18.3	12.4
10月	40.7	9.5	1.6	26.8	20.2	17.3	11.0
11月	39.3	9.2	1.7	26.4	19.5	16.1	10.0
12月	39.2	8.6	1.5	26.5	18.6	15.5	7.8

2016年	色	美	華やか・豪華	可愛らしさ	視覚に関する用語	季節感	場・機会
1月	41.9	8.8	1.6	26.1	19.2	16.1	7.8
2月	45.2	10.0	1.8	27.1	21.5	18.1	9.8
3月	45.8	11.0	1.8	27.4	23.2	20.0	10.7
4月	47.6	10.4	1.2	25.9	24.1	20.8	8.9
5月	50.9	10.4	1.4	26.6	26.4	21.6	8.7
6月	52.3	9.6	1.3	28.5	27.5	21.2	8.4
7月	50.8	9.1	1.2	33.0	28.0	20.6	9.2
8月	48.9	9.1	1.3	33.3	27.4	19.6	10.9
9月	45.6	10.1	1.7	27.0	24.1	21.9	10.1
10月	46.3	10.4	1.5	25.5	23.6	20.9	8.9
11月	45.6	9.5	1.6	25.1	22.4	20.5	7.5
12月	45.0	9.0	1.5	25.9	22.0	18.0	6.6

2017年	色	美	華やか・豪華	可愛らしさ	視覚に関する用語	季節感	場・機会
1月	46.7	9.0	1.7	26.1	20.8	18.4	7.0
2月	49.3	10.6	1.7	26.2	23.5	20.6	8.8
3月	49.9	11.9	1.9	27.2	25.1	22.1	9.7
4月	52.1	11.0	1.5	27.4	27.3	22.6	9.6
5月	49.4	10.2	1.5	28.0	27.6	22.2	8.4
6月	50.8	9.5	1.3	29.3	28.3	20.9	8.2
7月	47.2	8.8	1.3	32.6	27.8	21.8	8.7
8月	45.3	8.7	1.3	33.2	26.8	19.7	10.0
9月	46.7	10.1	1.5	29.9	25.4	21.8	9.1
10月	46.5	10.0	1.6	27.1	23.5	22.0	8.8
11月	46.0	9.9	1.5	26.4	23.2	19.9	7.5
12月	45.6	9.0	1.2	25.2	22.3	18.4	5.9

2018年	色	美	華やか・豪華	可愛らしさ	視覚に関する用語	季節感	場・機会
1月	45.8	8.9	1.6	24.7	20.9	18.2	7.4
2月	47.1	10.5	1.6	25.6	22.9	21.1	8.2
3月	47.4	10.9	1.6	25.6	24.8	21.9	9.1
4月	49.5	10.3	1.3	25.9	26.5	23.0	8.1
5月	49.4	9.4	1.3	26.6	27.9	22.4	7.7
6月	50.1	9.2	1.1	28.1	29.2	21.4	7.6
7月	49.1	8.8	1.2	29.5	28.4	24.0	8.0
8月	49.9	8.8	1.1	29.8	28.3	22.4	8.8
9月	47.1	9.2	1.3	27.2	24.9	23.2	8.5
10月	46.7	9.6	1.1	26.2	24.1	21.3	7.5
11月	47.7	9.5	1.4	25.5	23.1	19.8	6.6
12月	46.7	9.1	1.0	25.0	23.2	18.6	5.5

2019年	色	美	華やか・豪華	可愛らしさ	視覚に関する用語	季節感	場・機会
1月	48.5	9.4	1.4	24.7	22.1	17.3	6.7
2月	46.6	10.2	1.4	25.1	23.4	19.8	8.3
3月	49.0	11.1	1.4	25.8	25.8	21.4	8.9
4月	50.5	10.7	1.2	26.5	27.9	23.5	8.0
5月	49.6	9.9	1.0	25.6	29.0	23.7	6.9
6月	51.3	9.6	1.1	26.8	30.2	22.2	7.2
7月	49.2	9.5	1.2	28.2	29.8	20.6	8.7
8月	47.5	8.9	1.3	28.8	28.5	24.1	10.6
9月	46.1	10.1	1.4	27.2	26.8	23.5	8.0
10月	47.7	10.1	1.3	26.4	26.4	21.8	8.1
11月	48.1	10.1	1.3	25.5	25.7	20.5	7.1
12月	48.7	9.3	1.2	23.9	24.4	19.2	5.9

3. 「おしゃれ」に関する類似度

3では語句の類似度を用いて、「おしゃれ」に関して定量的な分析を行う。「おしゃれ」という抽象的な語句について、他のどのような語句と類似度が高いかを分析することで、「おしゃれ」に関するイメージを具体化したい。

2で行った形態素解析の結果から、語句の数値化（ベクトル）を行い、語句間のコサイン類似度を計算できる。Word2Vecは、文章中の単語を数値ベクトルに変換してその意味を把握する自然言語処理の手法である。Pythonを用いて実装できる。Word2Vecを用いるには既存の辞書を使う方法と新たに辞書を作成する方法があるが、本稿ではレディースファッションに関するクチコミに焦点をあてているため、レディースファッションに関するクチコミのみを用いて新たに辞書を作成した。辞書の作成には学習データ数もかなり必要なので、年単位クチコミに関して語句間のコサイン類似度を計算した。

図表2がその結果である。各年の「おしゃれ」という語句と類似度が相対的に高い語句を探索した。2で取り上げたカテゴリーから「場・機会」を外し、「季節感」を時期に変えたものに感情表現や手ざわりに関係する語句をもとにして、各年の分析のなかから候補を選んだ。「似合う」という語句はどの年次もプラスの値であり、比較的安定して「おしゃれ」と類似度が高かった。年別にみると、2015年はフランネル、似合う、ネイビー、カッコイイ、ステキが高い類似度になった。2016年はステキ、コーディネート、フランネル、グレー、ラベンダー、合うといった語句が相対的に高い値となった。2017年はかっこいい、モスグリーン、合う、似合う、カッコイイが相対的に高い値となった。2018年は化粧、写真、時期、かっこいい、びっくりが相対的に高い値となった。2019年は手ざわり、大事、感謝、サプライズ、信頼が高い値となった。2019年は感情や気持ちに関する語句が「おしゃれ」と高い類似度となっている。この傾向が2019年に特有なものか、それ以降の

図表2 「おしゃれ」に関するコサイン類似度

	2015	2016	2017	2018	2019
感激	0.0331	-0.2263	-0.4251	0.1825	0.2551
大事	0.0415	0.1491	-0.0944	0.1073	0.6868
感謝	0.0903	0.0257	0.2928	0.0660	0.5096
信頼	-0.1300	-0.1197	0.0195	0.0552	0.3625
サプライズ	0.1573	-0.0169	0.0773	0.0901	0.4298
びっくり	0.2639	-0.0523	0.2492	0.2879	-0.0549
似合う	0.5638	0.2189	0.3324	0.0750	0.1778
合う	0.2890	0.2846	0.3354	0.0393	-0.2076
組み合わせる	-0.0919	0.0443	0.0788	0.1349	0.1159
コーディネート	0.1589	0.3978	0.1986	-0.1682	-0.2387
アレンジ	-0.0765	-0.1068	-0.0498	0.1620	0.0103
着こなせる	-0.1417	0.0133	0.1241	-0.0420	0.1304
センス	-0.2261	0.0251	-0.1726	0.0460	-0.2190
ステキ	0.3364	0.4278	-0.1624	-0.1108	-0.0778
かっこいい	-0.0243	-0.0245	0.7288	0.3060	-0.0171
カッコイイ	0.3478	-0.0280	0.2963	-0.1153	-0.0346
さわやか	0.0774	0.1868	0.1874	-0.0374	-0.0365
可愛らしい	0.1303	0.2694	-0.0184	-0.0945	-0.2700
きれい	0.1908	0.0829	0.0943	-0.0402	0.0011
華やか	0.1096	0.0578	0.1856	-0.1554	0.0538
ゴージャス	-0.2150	-0.1470	-0.1125	-0.1344	0.0181
シック	-0.0369	-0.0553	0.0506	0.1617	0.0784
セクシー	0.2458	0.1184	0.2435	-0.0936	-0.0732
映える	-0.1497	-0.1344	-0.0808	0.1353	0.0840
一目	0.2432	-0.1646	-0.0559	-0.2268	-0.1515
見える	0.0194	0.1024	-0.0804	0.3710	0.0248
写真	0.2241	0.0475	0.3109	0.0802	-0.1031
鏡	-0.1850	0.2685	0.2669	-0.1321	0.0456
シルエット	0.3204	-0.3616	0.1780	0.1470	0.1793
赤	0.0966	0.1890	0.2078	0.2542	0.0137
ピンク	0.2528	0.0147	0.2034	-0.0310	-0.1048
モスグリーン	0.1028	-0.1714	0.4801	0.1101	-0.2139
ネイビー	0.3843	0.0654	0.0780	0.1214	0.2415
黒	-0.1495	-0.0256	-0.3259	0.1237	-0.1122
ブラック	-0.1786	-0.2341	-0.2037	-0.2135	-0.0473
グレー	-0.1359	0.3384	-0.0770	-0.2392	0.1348
ベージュ	0.1739	0.1296	0.1146	-0.0125	-0.0622
パール	-0.3297	0.0018	0.2572	-0.0466	0.0621
ローズ	0.0998	0.0874	0.1305	-0.0087	0.0292
ラベンダー	0.1805	0.2873	0.0705	-0.0117	-0.0373
化粧品	0.0909	-0.0483	0.3213	0.4962	0.0933
手ざわり	0.1604	-0.1552	0.2099	0.1186	0.7022
フランネル	0.6804	0.3435	0.1230	0.1191	0.0964
カシミア	0.0645	-0.2153	0.2869	-0.0511	0.2295
シルク	0.0626	0.2643	-0.1754	0.1810	0.0538
時期	-0.0425	0.4143	0.2631	0.3378	0.1839

傾向になるのかは、2020年以降の分析で明らかとなる。また2017年に感激という語句が「おしゃれ」と相対的に強いマイナスの類似度となっていることは現時点で不明であり今後引き続き分析したい。

4. 評価ポイントに関する順序プロビット分析

4章と次の5章は、評価点に関わる分析である。クチコミ投稿者がつけた評価ポイントであり、元のデータでは0から5の6段階であったが、先述した理由により本稿の分析には1から5が対象となる。本章の分析では、評価ポイントを被説明変数とした統計分析の手法の一つである、順序プロビットモデルを用いる。評価ポイント、クチコミ以外に使い道、リピーターかどうか（頻度）というデータも利用できる。使い道はプレゼントかそれ以外に分けた（プレゼント：1，それ以外：0）。クチコミに関する変数も該当した場合を1，該当しない場合を0とした。2015年から2019年まで月単位で推計した。

図表3が推計結果である。すべての期間で、1%水準で統計的に有意であった説明変数は、使い道（プレゼント）、頻度、語句1（エレガント、ゴージャス、華やか、豪華）、語句3（可愛い、かわいい）、語句6（オシャレ、おしゃれ、お洒落）、語句8（デザイン）、語句17（生地）である。符号がマイナスであった変数は、語句8（デザイン）、語句17（生地）である。マイナスになった理由として考えられるのは、デザインや生地に関して不満があったクチコミであったのではないかと推察される。係数の大きさについては、語句1（エレガント、ゴージャス、華やか、豪華）、使い道、頻度といった変数が他と比較して大きく、評価ポイントに影響をもっていたことがわかる。

図表3 評価ポイントについての推計結果

説明変数/評価ポイント	計方法 Orderedprobit regression											
	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
使い道(プロセント:1, それ以外:0)	0.2548*** (0.0230)	0.2944*** (0.0271)	0.3186*** (0.0245)	0.3114*** (0.0244)	0.3576*** (0.0222)	0.3688*** (0.0231)	0.3501*** (0.0250)	0.2809*** (0.0247)	0.3052*** (0.0265)	0.3441*** (0.0273)	0.3388*** (0.0277)	0.3068*** (0.0232)
頻度(リポート:1, それ以外:0)	0.3361*** (0.0092)	0.3309*** (0.0103)	0.3352*** (0.0094)	0.3546*** (0.0101)	0.3376*** (0.0100)	0.3328*** (0.0091)	0.3385*** (0.0098)	0.3194*** (0.0105)	0.3529*** (0.0104)	0.3516*** (0.0100)	0.2973*** (0.0099)	0.3071*** (0.0089)
語句1("エレガント","ゴージャス","華やか","豪華",が含まれるケース:1, それ以外:0)	0.4135*** (0.0251)	0.3666*** (0.0245)	0.3930*** (0.0223)	0.4073*** (0.0274)	0.4123*** (0.0298)	0.4605*** (0.0289)	0.4374*** (0.0282)	0.3898*** (0.0322)	0.3712*** (0.0301)	0.4207*** (0.0290)	0.4507*** (0.0280)	0.5067*** (0.0321)
語句2("セクシー"が含まれるケース:1, それ以外:0)	0.1302** (0.0603)	0.0694 (0.0653)	0.1305* (0.0675)	0.2929*** (0.0681)	0.1300** (0.0674)	0.2703*** (0.0660)	0.1836*** (0.0650)	0.1769*** (0.0680)	0.1443** (0.0723)	0.2082*** (0.0737)	0.1030 (0.0737)	0.1503** (0.0758)
語句3("可愛い","かわいく"が含まれるケース:1, それ以外:0)	0.1438*** (0.0076)	0.1033*** (0.0086)	0.1206*** (0.0076)	0.1523*** (0.0083)	0.1353*** (0.0083)	0.1202*** (0.0072)	0.1400*** (0.0072)	0.1064*** (0.0077)	0.1305*** (0.0086)	0.1578*** (0.0085)	0.1508*** (0.0087)	0.1679*** (0.0088)
語句4("キレイ","きれい","美しい","美容","綺麗",が含まれるケース:1, それ以外:0)	0.3184*** (0.0105)	0.2929*** (0.0111)	0.2954*** (0.0097)	0.2769*** (0.0105)	0.2412*** (0.0107)	0.2804*** (0.0100)	0.2447*** (0.0105)	0.2553*** (0.0112)	0.2988*** (0.0113)	0.3157*** (0.0110)	0.3134*** (0.0114)	0.3141*** (0.0118)
語句5("カッコイイ","カっこいい"が含まれるケース:1, それ以外:0)	0.3971*** (0.0958)	0.2547** (0.1095)	0.3306*** (0.1127)	0.3112*** (0.1169)	0.3332*** (0.1168)	0.2377** (0.1025)	0.2732** (0.1071)	0.0871*** (0.1492)	0.5026*** (0.1264)	0.2489** (0.1196)	0.4401*** (0.1138)	0.2924** (0.1165)
語句6("オシャレ","おしゃれ","お洒落"が含まれるケース:1, それ以外:0)	0.3637*** (0.0270)	0.3063*** (0.0283)	0.3565*** (0.0276)	0.3731*** (0.0300)	0.3572*** (0.0288)	0.3412*** (0.0248)	0.3916*** (0.0283)	0.3468*** (0.0301)	0.3231*** (0.0282)	0.3700*** (0.0282)	0.3700*** (0.0285)	0.3587*** (0.0295)
語句7("ソフエッド",が含まれるケース:1, それ以外:0)	-0.0052 (0.0090)	-0.0317* (0.0191)	-0.0170 (0.0172)	-0.0186 (0.0187)	-0.0160 (0.0191)	0.0036 (0.0181)	-0.0104 (0.0195)	-0.0096 (0.0211)	-0.0096 (0.0211)	-0.0953 (0.0180)	-0.0169 (0.0179)	-0.0205 (0.0183)
語句8("デザイン",が含まれるケース:1, それ以外:0)	-0.1012*** (0.0090)	-0.0853*** (0.0100)	-0.1053*** (0.0089)	-0.1052*** (0.0089)	-0.1101*** (0.0099)	-0.1159*** (0.0088)	-0.1104*** (0.0090)	-0.1099*** (0.0095)	-0.1297*** (0.0101)	-0.1052*** (0.0089)	-0.1133*** (0.0100)	-0.1411*** (0.0104)
語句9("ベージュ",が含まれるケース:1, それ以外:0)	-0.0434*** (0.0163)	-0.0031 (0.0175)	-0.0236 (0.0149)	-0.0630*** (0.0171)	-0.0701*** (0.0198)	-0.0715*** (0.0178)	-0.0602*** (0.0203)	-0.0563*** (0.0215)	-0.0629*** (0.0201)	-0.0464** (0.0186)	-0.0655*** (0.0183)	-0.0207 (0.0199)
語句10("赤","レッド"が含まれるケース:1, それ以外:0)	-0.0328** (0.0205)	-0.0394 (0.0239)	-0.0295 (0.0246)	0.0544** (0.0275)	0.0285 (0.0277)	-0.0225 (0.0243)	0.0420* (0.0244)	0.0433* (0.0251)	0.0420* (0.0250)	0.0476** (0.0229)	0.0314 (0.0241)	0.0073 (0.0240)
語句11("ピンク",が含まれるケース:1, それ以外:0)	0.0263 (0.0195)	0.0107 (0.0206)	-0.0021 (0.0171)	-0.0178 (0.0178)	-0.0262 (0.0193)	-0.0447*** (0.0166)	0.0136 (0.0165)	0.0097 (0.0184)	-0.0070 (0.0209)	0.0109 (0.0222)	-0.0011 (0.0232)	0.0347 (0.0248)
語句12("イエロー","黄色"が含まれるケース:1, それ以外:0)	-0.1333*** (0.0346)	-0.1109*** (0.0355)	-0.1291*** (0.0277)	-0.1555*** (0.0287)	-0.1003*** (0.0246)	-0.0646*** (0.0224)	-0.0450** (0.0222)	-0.0661** (0.0257)	-0.0854** (0.0337)	-0.1296*** (0.0395)	-0.1618*** (0.0408)	-0.1987*** (0.0445)
語句13("グリーン","緑"が含まれるケース:1, それ以外:0)	0.0861*** (0.0204)	0.0309 (0.0231)	0.0251 (0.0207)	0.0441** (0.0220)	0.0376* (0.0218)	-0.0054 (0.019)	-0.0179 (0.0199)	-0.0034 (0.0208)	0.0148 (0.0226)	0.0085* (0.0221)	0.0607*** (0.0232)	0.0279 (0.0249)
語句14("ブルー",が含まれるケース:1, それ以外:0)	0.0650 (0.0142)	0.1005*** (0.0160)	0.0732*** (0.0147)	0.0600*** (0.0162)	0.1800*** (0.0170)	0.0813*** (0.0151)	0.0548*** (0.0169)	0.0660*** (0.0187)	0.0898*** (0.0154)	0.0668*** (0.0154)	0.0771*** (0.0154)	0.0551 (0.0152)
語句15("黒","ブラック"が含まれるケース:1, それ以外:0)	0.0534*** (0.0100)	0.0671*** (0.0111)	0.0981*** (0.0102)	0.0606*** (0.0115)	0.0774*** (0.0114)	0.0974*** (0.0103)	0.0855*** (0.0106)	0.0692*** (0.0113)	0.1047*** (0.0114)	0.0919*** (0.0112)	0.0995*** (0.0119)	0.1087*** (0.0119)
語句16("肌ざわり","手触り"が含まれるケース:1, それ以外:0)	0.1747*** (0.0334)	0.1866*** (0.0391)	0.1680*** (0.0405)	0.2201*** (0.0427)	0.1636*** (0.0453)	0.2009*** (0.0448)	0.2989*** (0.0445)	0.1571*** (0.0477)	0.1708*** (0.0443)	0.1236*** (0.0372)	0.2155*** (0.0371)	0.1915*** (0.0354)
語句17("生地"が含まれるケース:1, それ以外:0)	-0.1601*** (0.0071)	-0.1634*** (0.0077)	-0.1592*** (0.0067)	-0.1371*** (0.0072)	-0.1311*** (0.0072)	-0.1243*** (0.0064)	-0.1300*** (0.0069)	-0.1241*** (0.0074)	-0.1376*** (0.0074)	-0.1278*** (0.0074)	-0.1618*** (0.0076)	-0.1806*** (0.0080)
サンプル数	166,204	135,838	171,848	147,200	145,751	179,984	174,098	151,328	139,212	142,076	138,785	130,576
Wald χ^2 square	3,946.60	2,960.10	4,014.97	3,316.54	3,148.05	3,829.91	3,374.94	2,623.08	3,186.38	3,397.47	3,290.26	3,441.55
Prob > χ^2 square	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
擬似決定係数	0.0092	0.0088	0.0093	0.0089	0.0087	0.0087	0.0080	0.0071	0.0092	0.0093	0.0091	0.0101

***, **, * < 0.05, * < 0.1以下同じ

レディース・ファッションに関するインターネット上のクチコミについてのデータ解析 (峰竜)

2016年

被説明変数評価ポイント

相計方法 Orderedprobit regression

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
使い道 (ブレゼント: 1, それ以外: 0)	0.3515 *** (0.0310)	0.2686 *** (0.0389)	0.3228 *** (0.0400)	0.3342 *** (0.0381)	0.3586 *** (0.0433)	0.4609 *** (0.0408)	0.3900 *** (0.0413)	0.2102 *** (0.0446)	0.2493 *** (0.0428)	0.3062 *** (0.0438)	0.3888 *** (0.0391)	0.3923 *** (0.0331)
頻度 (リピーター: 1, それ以外: 0)	0.2839 *** (0.0122)	0.2839 *** (0.0137)	0.2924 *** (0.0137)	0.3004 *** (0.0131)	0.3213 *** (0.0140)	0.3306 *** (0.0137)	0.3565 *** (0.0146)	0.3221 *** (0.0170)	0.3314 *** (0.0143)	0.3360 *** (0.0143)	0.3260 *** (0.0134)	0.3423 *** (0.0128)
語句1 ("エレガント", "ゴージャズ", "華やか", "豪華", が含まれるケース: 1, それ以外: 0)	0.4350 *** (0.0371)	0.3877 *** (0.0401)	0.3186 *** (0.0388)	0.3166 *** (0.0416)	0.5144 *** (0.0512)	0.4619 *** (0.0492)	0.3406 *** (0.0536)	0.5780 *** (0.0597)	0.5293 *** (0.0496)	0.5391 *** (0.0487)	0.6574 *** (0.0456)	0.4881 *** (0.0429)
語句2 ("可愛い", "かわいい") が含まれるケース: 1, それ以外: 0)	0.2161 * (0.1135)	0.2130 ** (0.0887)	0.1929 (0.1003)	0.0522 (0.0829)	0.4710 *** (0.1121)	0.1461 (0.0831)	0.1405 * (0.0849)	0.1619 * (0.0832)	0.2536 ** (0.1034)	0.2193 * (0.1129)	0.4217 *** (0.1044)	0.2631 *** (0.0668)
語句3 ("可愛い", "かわいい") が含まれるケース: 1, それ以外: 0)	0.1687 *** (0.0111)	0.1950 *** (0.0128)	0.1539 *** (0.0123)	0.1732 *** (0.0119)	0.1769 *** (0.0135)	0.1804 *** (0.0119)	0.2043 *** (0.0117)	0.2073 *** (0.0134)	0.1973 *** (0.0135)	0.1633 *** (0.0133)	0.1891 *** (0.0120)	0.1751 *** (0.0120)
語句4 ("キレイ", "きれい", "美しい", "美容", "綺麗", が含まれるケース: 1, それ以外: 0)	0.3404 *** (0.0149)	0.2908 ** (0.0161)	0.2955 *** (0.0151)	0.3235 *** (0.0147)	0.2683 *** (0.0168)	0.2839 *** (0.0160)	0.2730 *** (0.0164)	0.3027 *** (0.0193)	0.2816 *** (0.0170)	0.3163 *** (0.0167)	0.3175 *** (0.0159)	0.3564 *** (0.0159)
語句5 ("カッコイイ", "カッコイイ") が含まれるケース: 1, それ以外: 0)	0.8707 *** (0.1636)	0.2897 * (0.1767)	0.5704 *** (0.1468)	0.3340 * (0.1724)	0.7851 *** (0.1907)	0.2930 ** (0.1418)	0.2721 (0.1791)	0.0249 (0.2388)	0.1847 (0.1481)	0.3059 (0.1981)	0.5746 *** (0.1678)	0.5155 *** (0.1381)
語句6 ("オシャレ", "おしゃれ", "お洒落") が含まれるケース: 1, それ以外: 0)	0.3941 *** (0.0361)	0.3488 *** (0.0366)	0.3201 *** (0.0415)	0.3132 *** (0.0382)	0.3539 *** (0.0414)	0.3825 *** (0.0382)	0.3303 *** (0.0415)	0.3364 *** (0.0414)	0.3190 *** (0.0407)	0.3857 *** (0.0391)	0.4104 *** (0.0375)	0.4158 *** (0.0360)
語句7 ("シルエット", が含まれるケース: 1, それ以外: 0)	-0.0207 (0.0239)	-0.0405 (0.0259)	-0.1810 (0.0259)	-0.0076 (0.0248)	-0.0171 (0.0281)	-0.0542 ** (0.0252)	-0.0380 (0.0279)	-0.1600 (0.0327)	-0.0068 (0.0268)	-0.0204 (0.0254)	0.0039 (0.0239)	-0.0096 (0.0227)
語句8 ("デザイン", が含まれるケース: 1, それ以外: 0)	-0.1258 *** (0.0126)	-0.1267 *** (0.0141)	-0.1216 *** (0.0139)	-0.1234 *** (0.0137)	-0.1137 *** (0.0154)	-0.1438 *** (0.0137)	-0.1371 *** (0.0140)	-0.0938 *** (0.0159)	-0.0858 *** (0.0152)	-0.0655 *** (0.0149)	-0.0797 *** (0.0141)	-0.1283 *** (0.0137)
語句9 ("ベージュ", が含まれるケース: 1, それ以外: 0)	-0.0431 * (0.0237)	-0.0289 (0.0251)	-0.0492 ** (0.0230)	-0.0418 * (0.0231)	-0.0567 ** (0.0282)	-0.0205 (0.0264)	-0.0076 (0.0295)	0.0026 (0.0354)	-0.0469 (0.0286)	-0.0144 (0.0279)	-0.0577 ** (0.0254)	0.0241 (0.0265)
語句10 ("赤", "レッド" が含まれるケース: 1, それ以外: 0)	-0.0289 (0.0291)	0.0072 (0.0363)	-0.0249 (0.0386)	0.0896 ** (0.0402)	-0.1389 *** (0.0458)	-0.0019 (0.0384)	-0.0115 (0.0385)	0.0842 ** (0.0414)	-0.0131 (0.0395)	-0.0522 (0.0347)	-0.0026 (0.0323)	0.0003 (0.0303)
語句11 ("ピンク", が含まれるケース: 1, それ以外: 0)	0.0079 (0.0288)	-0.0011 (0.0315)	-0.0254 (0.0268)	-0.0371 (0.0259)	-0.0790 *** (0.0288)	-0.0435 * (0.0234)	0.0258 (0.0268)	0.0255 (0.0325)	-0.0125 (0.0333)	-0.0657 (0.0347)	-0.0182 (0.0326)	0.0252 (0.0311)
語句12 ("イエロー", "黄色" が含まれるケース: 1, それ以外: 0)	-0.1465 *** (0.0518)	-0.0612 (0.0549)	-0.1115 ** (0.0480)	-0.0654 ** (0.0429)	-0.1266 *** (0.0451)	-0.0428 (0.0391)	-0.0518 (0.0400)	-0.1071 ** (0.0462)	-0.0907 (0.0375)	-0.0601 (0.0608)	-0.1847 *** (0.0655)	-0.2328 *** (0.0625)
語句13 ("グリーン", "緑" が含まれるケース: 1, それ以外: 0)	0.0221 (0.0307)	0.0578 * (0.0348)	0.0666 * (0.0344)	0.0402 (0.0327)	0.0496 (0.0352)	0.0374 (0.0318)	0.0606 * (0.0318)	0.0293 (0.0370)	0.0633 * (0.0374)	0.0331 (0.0349)	0.1124 *** (0.0359)	0.0538 (0.0332)
語句14 ("ブルー", が含まれるケース: 1, それ以外: 0)	0.0400 ** (0.0185)	0.0739 *** (0.0218)	0.0226 (0.0221)	0.0560 *** (0.0202)	0.0692 *** (0.0236)	0.0838 *** (0.0215)	0.0507 ** (0.0236)	0.0433 (0.0285)	0.0437 * (0.0246)	0.0600 *** (0.0218)	0.0290 (0.0199)	0.0892 (0.0190)
語句15 ("黒", "ブラック" が含まれるケース: 1, それ以外: 0)	0.0625 *** (0.0144)	0.0685 *** (0.0161)	0.0962 *** (0.0162)	0.0712 *** (0.0155)	0.1009 *** (0.0173)	0.0719 *** (0.0154)	0.1064 *** (0.0158)	0.0790 *** (0.0181)	0.0867 *** (0.0169)	0.1030 *** (0.0164)	0.0892 *** (0.0153)	0.0447 *** (0.0152)
語句16 ("肌ざわり", "手触り" が含まれるケース: 1, それ以外: 0)	-0.1716 *** (0.0101)	-0.1624 *** (0.0111)	-0.1535 *** (0.0106)	-0.1384 *** (0.0088)	-0.1267 *** (0.0111)	-0.1211 *** (0.0101)	-0.1438 *** (0.0095)	-0.1509 *** (0.0125)	-0.1674 *** (0.0119)	-0.1454 *** (0.0112)	-0.1564 *** (0.0109)	-0.1804 *** (0.0086)
サンプル数	83,640	64,945	67,797	74,353	55,817	68,279	64,970	48,464	56,538	58,586	65,532	69,901
Wald χ^2 square	2,083.07	1,585.13	1,612.85	1,907.01	1,429.53	1,742.74	1,686.41	1,212.54	1,518.86	1,545.24	1,961.48	2,145.08
Prob > χ^2 square	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
擬似決定係数	0.0065	0.0094	0.0085	0.0087	0.0087	0.0089	0.0100	0.0084	0.0101	0.0089	0.0108	0.0112

2017年

報説明変数評価ポイント

相計方法 Orderedprobit regression

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
使い道 (ブレゼント: 1, それ以外: 0)	0.3824 *** (0.0386)	0.2390 *** (0.0490)	0.3602 *** (0.0456)	0.3990 *** (0.0536)	0.3458 *** (0.0440)	0.2976 *** (0.0418)	0.3525 *** (0.0401)	0.4374 *** (0.0436)	0.3337 *** (0.0489)	0.5094 *** (0.0534)	0.3167 *** (0.0444)	0.4069 *** (0.0430)
頻度 (リビート: 1, それ以外: 0)	0.0139 *** (0.0139)	0.0163 *** (0.0163)	0.0148 *** (0.0148)	0.0162 *** (0.0162)	0.0156 *** (0.0156)	0.0144 *** (0.0144)	0.0148 *** (0.0148)	0.0155 *** (0.0155)	0.0161 *** (0.0161)	0.0149 *** (0.0149)	0.0140 *** (0.0140)	0.0138 *** (0.0138)
語句1 ("エレガント", "ゴージャズ", "華やか", "豪華", が含まれるケース: 1, それ以外: 0)	0.5297 *** (0.0472)	0.4922 *** (0.0522)	0.5285 *** (0.0433)	0.5398 *** (0.0520)	0.6539 *** (0.0562)	0.413 *** (0.0511)	0.5451 *** (0.0522)	0.3992 *** (0.0563)	0.4363 *** (0.0549)	0.4246 *** (0.0484)	0.5366 *** (0.0487)	0.5735 *** (0.0511)
語句2 ("可愛い", "かわいい") が含まれるケース: 1, それ以外: 0)	0.2177 ** (0.1006)	0.2407 ** (0.1215)	0.1969 ** (0.0884)	0.3993 *** (0.1199)	0.3555 *** (0.1220)	0.3777 *** (0.1808)	0.1012 (0.1012)	0.4038 *** (0.0838)	0.3807 *** (0.1155)	0.0236 (0.1063)	0.2030 * (0.1095)	0.3167 *** (0.1189)
語句3 ("可愛い", "かわいい") が含まれるケース: 1, それ以外: 0)	0.1522 *** (0.0134)	0.1824 *** (0.0157)	0.1680 *** (0.0136)	0.1575 *** (0.0147)	0.1646 *** (0.0144)	0.1787 *** (0.0131)	0.1677 *** (0.0127)	0.1770 *** (0.0131)	0.1718 *** (0.0146)	0.1520 *** (0.0137)	0.1833 *** (0.0133)	0.1762 *** (0.0126)
語句4 ("キレイ", "きれい", "美しい", "美容", "綺麗", が含まれるケース: 1, それ以外: 0)	0.3634 *** (0.0175)	0.2893 *** (0.0191)	0.3435 *** (0.0163)	0.3333 *** (0.0183)	0.3112 *** (0.0188)	0.3101 *** (0.0180)	0.2896 *** (0.0184)	0.2737 *** (0.0192)	0.3122 *** (0.0193)	0.2811 *** (0.0175)	0.3611 *** (0.0170)	0.3669 *** (0.0167)
語句5 ("カッコイイ", "カッコイイ") が含まれるケース: 1, それ以外: 0)	0.7241 *** (0.1987)	0.7139 *** (0.2105)	0.5309 *** (0.1701)	0.2166 (0.1396)	0.1034 (0.1670)	0.6291 *** (0.2112)	1.0018 *** (0.2595)	0.2450 (0.1780)	1.1178 *** (0.2847)	0.6562 *** (0.1918)	0.5138 ** (0.2151)	0.6000 *** (0.1792)
語句6 ("オシャレ", "おしゃれ", "お洒落") が含まれるケース: 1, それ以外: 0)	0.3331 *** (0.0403)	0.3443 *** (0.0478)	0.3805 *** (0.0412)	0.3995 *** (0.0485)	0.4096 *** (0.0455)	0.4635 *** (0.0445)	0.4039 *** (0.0431)	0.4142 *** (0.0445)	0.4687 *** (0.0482)	0.3452 *** (0.0404)	0.4389 *** (0.0418)	0.3544 *** (0.0405)
語句7 ("セクシー", "セクシー") が含まれるケース: 1, それ以外: 0)	0.030 (0.0266)	-0.0178 (0.0288)	-0.0069 (0.0269)	0.0246 (0.0298)	0.0206 (0.0303)	-0.0500 * (0.0292)	0.0209 (0.0317)	-0.0402 (0.0318)	-0.0444 (0.0301)	-0.0312 (0.0257)	0.0316 (0.0238)	-0.0189 (0.0234)
語句8 ("デザイン", "デザイン") が含まれるケース: 1, それ以外: 0)	-0.1287 *** (0.0149)	-0.1084 *** (0.0171)	-0.1035 *** (0.0151)	-0.1125 *** (0.0167)	-0.1133 *** (0.0163)	-0.1093 *** (0.0152)	-0.1085 *** (0.0144)	-0.1362 *** (0.0152)	-0.1023 *** (0.0168)	-0.1106 *** (0.0152)	-0.1196 *** (0.0148)	-0.0856 *** (0.0147)
語句9 ("ベージュ", "ベージュ") が含まれるケース: 1, それ以外: 0)	-0.0672 ** (0.0283)	-0.0331 (0.0290)	-0.0069 (0.0249)	-0.0184 (0.0281)	-0.0223 (0.0297)	-0.0600 ** (0.0301)	-0.0217 (0.0323)	-0.0642 ** (0.0326)	-0.0511 (0.0311)	-0.0125 (0.0287)	0.0386 (0.0269)	0.0413 (0.0275)
語句10 ("赤", "レッド") が含まれるケース: 1, それ以外: 0)	-0.0324 (0.0361)	-0.0725 * (0.0421)	0.1075 *** (0.0411)	0.1049 ** (0.0461)	0.0411 (0.0452)	-0.0286 (0.0373)	0.0138 (0.0398)	-0.0132 (0.0413)	0.0215 (0.0403)	-0.0681 (0.0359)	-0.0232 (0.0336)	0.0086 (0.0329)
語句11 ("ピンク", "ピンク") が含まれるケース: 1, それ以外: 0)	0.0108 (0.0318)	0.0171 (0.0310)	-0.0379 (0.0260)	-0.0457 (0.0285)	-0.0272 (0.0301)	-0.0188 (0.0286)	-0.0386 (0.0298)	-0.1165 *** (0.0312)	-0.0792 ** (0.0358)	0.0229 (0.0350)	-0.0082 (0.0334)	0.0017 (0.0324)
語句12 ("イエロー", "黄色") が含まれるケース: 1, それ以外: 0)	-0.1944 *** (0.0662)	-0.0096 (0.0668)	-0.1793 *** (0.0480)	-0.1288 *** (0.0445)	0.0045 (0.0438)	-0.0651 * (0.0392)	0.0005 (0.0439)	-0.0742 (0.0462)	-0.0699 (0.0564)	-0.1318 ** (0.0563)	-0.0985 (0.0628)	-0.1005 * (0.0591)
語句13 ("グリーン", "緑") が含まれるケース: 1, それ以外: 0)	-0.0233 (0.0382)	0.0832 * (0.0485)	0.0173 (0.0370)	-0.0067 (0.0397)	0.0715 * (0.0384)	-0.0480 (0.0447)	0.0145 (0.0319)	0.0145 (0.0381)	0.0532 (0.0388)	0.0555 (0.0343)	0.0697 ** (0.0340)	0.0223 (0.0345)
語句14 ("ブルー", "ブルー") が含まれるケース: 1, それ以外: 0)	0.0668 *** (0.0206)	0.0387 (0.0247)	0.0724 *** (0.0232)	0.0792 *** (0.0265)	0.0873 *** (0.0267)	0.0401 (0.0252)	0.0121 (0.0274)	0.0411 (0.0317)	0.0592 ** (0.0283)	0.0406 * (0.0285)	0.0607 *** (0.0214)	0.0859 *** (0.0213)
語句15 ("黒", "ブラック") が含まれるケース: 1, それ以外: 0)	-0.0076 (0.0161)	0.0249 * (0.0189)	0.0597 *** (0.0172)	0.0228 (0.0188)	0.1116 *** (0.0183)	0.0783 *** (0.0180)	0.0850 *** (0.0177)	0.1045 *** (0.0185)	0.0983 *** (0.0191)	0.1117 *** (0.0177)	0.0878 *** (0.0161)	0.0376 ** (0.0154)
語句16 ("肌ざわり", "手触り") が含まれるケース: 1, それ以外: 0)	0.0847 * (0.0490)	0.0906 (0.0638)	0.1982 *** (0.0604)	0.2146 *** (0.0722)	0.1523 ** (0.0702)	0.0616 (0.0713)	0.1150 (0.0715)	0.1273 * (0.0669)	0.1520 ** (0.0700)	0.0242 *** (0.0587)	0.1415 *** (0.0503)	0.1453 *** (0.0494)
語句17 ("生地") が含まれるケース: 1, それ以外: 0)	-0.1738 *** (0.0121)	-0.1774 *** (0.0155)	-0.1766 *** (0.0119)	-0.1908 *** (0.0123)	-0.1508 *** (0.0120)	-0.1352 *** (0.0115)	-0.1339 *** (0.0112)	-0.1348 *** (0.0119)	-0.1517 *** (0.0127)	-0.1567 *** (0.0114)	-0.1506 *** (0.0114)	-0.1790 *** (0.0111)
サンプル数	1,687.44	1,071.38	1,639.23	1,266.27	1,380.32	1,547.62	1,450.62	1,299.98	1,242.15	1,418.77	1,669.27	1,871.80
Wald χ^2 square	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Prob > χ^2 square	0.1068	0.0096	0.0110	0.0104	0.0107	0.0106	0.0096	0.0095	0.0106	0.0101	0.0105	0.0107
擬似決定係数												

レディース・ファッションに関するインターネット上のクチコミについてのデータ解析 (峰竜)

相計方法 Orderedprobit regression

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
使い道 (ブレゼント: 1, それ以外: 0)	0.3812 *** (0.0570)	0.3327 *** (0.0635)	0.4419 *** (0.0578)	0.3832 *** (0.0591)	0.3406 *** (0.0532)	0.3098 *** (0.0573)	0.5401 *** (0.0603)	0.4909 *** (0.0629)	0.3032 *** (0.0704)	0.5364 *** (0.0633)	0.4609 *** (0.0627)	0.4207 *** (0.0507)
頻度 (リピーター: 1, それ以外: 0)	0.3329 *** (0.0149)	0.2938 *** (0.0165)	0.3203 *** (0.0156)	0.341 *** (0.0160)	0.3470 *** (0.0155)	0.3569 *** (0.0148)	0.3546 *** (0.0149)	0.3363 *** (0.0168)	0.3461 *** (0.0170)	0.3366 *** (0.0158)	0.3125 *** (0.0157)	0.3013 *** (0.0145)
語句1 ("エレガント", "ゴージャズ", "華やか", "豪華", が含まれるケース: 1, それ以外: 0)	0.6439 *** (0.0526)	0.5092 *** (0.0534)	0.5148 *** (0.0465)	0.5710 *** (0.0531)	0.5647 *** (0.0535)	0.5068 *** (0.0588)	0.4767 *** (0.0561)	0.4902 *** (0.0612)	0.5370 *** (0.0612)	0.4581 *** (0.0568)	0.570 *** (0.0541)	0.5857 *** (0.0558)
語句2 ("セクシー"が含まれるケース: 1, それ以外: 0)	0.1790 (0.1135)	0.1911 (0.1418)	0.3303 *** (0.1215)	0.2708 *** (0.1042)	0.4469 *** (0.1253)	0.3390 *** (0.1090)	0.5229 *** (0.1215)	0.4092 *** (0.1201)	0.3178 *** (0.1235)	0.4516 *** (0.1885)	0.2433 * (0.1447)	0.2050 * (0.1091)
語句3 ("可愛い", "かわいい")が含まれるケース: 1, それ以外: 0)	0.1988 (0.0141)	0.1618 (0.0156)	0.1659 *** (0.0138)	0.1837 *** (0.0144)	0.1930 *** (0.0143)	0.1918 *** (0.0132)	0.1903 *** (0.0127)	0.2240 *** (0.0144)	0.2005 *** (0.0153)	0.2026 *** (0.0141)	0.2114 *** (0.0140)	0.2200 *** (0.0140)
語句4 ("キレイ", "きれい", "美しい", "美容", "綺麗", が含まれるケース: 1, それ以外: 0)	0.3758 *** (0.0185)	0.2898 *** (0.0186)	0.3292 *** (0.0170)	0.3429 *** (0.0181)	0.3239 *** (0.0190)	0.3352 *** (0.0182)	0.2779 *** (0.0177)	0.3203 *** (0.0204)	0.3402 *** (0.0206)	0.3092 *** (0.0184)	0.3472 *** (0.0182)	0.3900 *** (0.0172)
語句5 ("カッコイイ", "カッコイイ")が含まれるケース: 1, それ以外: 0)	0.4141 ** (0.2024)	0.5333 ** (0.2089)	0.1819 (0.1487)	0.3394 * (0.1851)	0.5438 ** (0.2262)	0.6191 ** (0.2508)	0.5075 ** (0.2381)	0.6107 ** (0.2685)	0.0890 (0.2016)	0.4023 (0.1717)	0.2617 * (0.1580)	0.5672 *** (0.1641)
語句6 ("オシャレ", "おしゃれ", "お洒落"が含まれるケース: 1, それ以外: 0)	0.3101 *** (0.0438)	0.3307 *** (0.0476)	0.2555 *** (0.0469)	0.3875 *** (0.0445)	0.3996 *** (0.0426)	0.4113 *** (0.0412)	0.4592 *** (0.0425)	0.4206 *** (0.0446)	0.2702 *** (0.0463)	0.3279 *** (0.0424)	0.4321 *** (0.0416)	0.3404 *** (0.0387)
語句7 ("シルエット", が含まれるケース: 1, それ以外: 0)	-0.0248 (0.0266)	-0.0344 (0.0289)	-0.0741 *** (0.0259)	-0.0156 (0.0287)	-0.0412 (0.0282)	-0.0347 (0.0279)	-0.0485 * (0.0281)	-0.091 (0.0302)	-0.0240 (0.0304)	-0.0499 * (0.0267)	-0.0222 (0.0248)	-0.0220 (0.0231)
語句8 ("デザイン", が含まれるケース: 1, それ以外: 0)	-0.0989 *** (0.0158)	-0.0911 *** (0.0174)	-0.1513 *** (0.0155)	-0.1467 *** (0.0165)	-0.1178 *** (0.0162)	-0.1215 *** (0.0153)	-0.1213 *** (0.0146)	-0.1345 *** (0.0165)	-0.0923 *** (0.0175)	-0.0914 *** (0.0162)	-0.0971 *** (0.0164)	-0.0813 *** (0.0151)
語句9 ("ベージュ", が含まれるケース: 1, それ以外: 0)	0.0105 (0.0286)	0.0223 (0.0296)	0.0358 (0.0247)	0.0268 (0.0272)	0.0388 (0.0246)	-0.0888 *** (0.0249)	-0.1836 *** (0.0278)	-0.2244 *** (0.0302)	-0.0521 * (0.0301)	-0.0035 (0.0278)	-0.0463 * (0.0275)	-0.0794 *** (0.0265)
語句10 ("赤", "レッド"が含まれるケース: 1, それ以外: 0)	-0.0529 (0.0378)	-0.0185 (0.0427)	-0.0418 (0.0414)	0.0099 (0.0471)	0.1231 *** (0.0439)	0.0485 (0.0401)	0.0523 (0.0383)	0.0243 (0.0411)	-0.019 (0.0418)	-0.0276 (0.0365)	-0.0249 (0.0358)	-0.0015 (0.0336)
語句11 ("ピンク", が含まれるケース: 1, それ以外: 0)	-0.0090 (0.0340)	-0.0394 (0.0338)	0.0142 (0.0283)	-0.0696 *** (0.0286)	0.0005 (0.0324)	-0.0668 *** (0.0312)	-0.0424 (0.0315)	-0.0109 (0.0380)	0.0038 (0.0393)	-0.0658 (0.0376)	0.0049 (0.0362)	0.0268 (0.0339)
語句12 ("イエロー", "黄色"が含まれるケース: 1, それ以外: 0)	-0.2735 *** (0.0591)	-0.1205 ** (0.0552)	-0.1285 *** (0.0439)	-0.0690 ** (0.0452)	-0.0281 (0.0440)	-0.0837 ** (0.0398)	-0.0496 (0.0392)	-0.3110 *** (0.0446)	-0.1483 *** (0.0545)	-0.0654 (0.0641)	-0.1364 ** (0.0637)	-0.1426 ** (0.0592)
語句13 ("グリーン", "緑"が含まれるケース: 1, それ以外: 0)	0.0499 (0.0385)	0.0010 (0.0495)	-0.0198 (0.0386)	-0.0578 (0.0375)	0.0071 (0.0337)	0.0256 (0.0325)	0.0261 (0.0315)	0.0838 ** (0.0389)	0.1119 *** (0.0390)	0.0594 * (0.0358)	-0.0016 (0.0363)	0.1471 *** (0.0363)
語句14 ("ブルー", が含まれるケース: 1, それ以外: 0)	0.0404 * (0.0220)	0.0280 (0.0247)	-0.0003 (0.0235)	0.0741 *** (0.0250)	0.0582 ** (0.0238)	0.0836 *** (0.0249)	0.0350 (0.0268)	0.0719 ** (0.0283)	0.0284 (0.0280)	0.0468 ** (0.0231)	0.1074 *** (0.0244)	0.0773 *** (0.0241)
語句15 ("黒", "ブラック"が含まれるケース: 1, それ以外: 0)	0.0165 (0.0170)	0.0512 *** (0.0192)	0.0857 *** (0.0180)	0.0462 ** (0.0184)	0.0469 *** (0.0180)	0.1089 *** (0.0170)	0.1427 *** (0.0166)	0.0909 *** (0.0186)	0.0925 *** (0.0191)	0.0993 *** (0.0171)	0.0891 *** (0.0167)	0.0645 *** (0.0158)
語句16 ("肌ざわり", "手触り"が含まれるケース: 1, それ以外: 0)	0.1817 *** (0.0540)	0.0834 (0.0598)	0.3315 *** (0.0654)	0.3081 *** (0.0688)	0.0452 (0.0751)	0.1580 ** (0.0643)	0.1574 ** (0.0692)	0.2665 *** (0.0827)	0.0804 (0.0680)	0.2070 *** (0.0574)	0.1172 ** (0.0544)	0.1930 *** (0.0516)
語句17 ("生地"が含まれるケース: 1, それ以外: 0)	-0.1285 *** (0.0125)	-0.1683 *** (0.0117)	-0.1405 *** (0.0116)	-0.1327 *** (0.0117)	-0.1290 *** (0.0117)	-0.1406 *** (0.0110)	-0.1446 *** (0.0108)	-0.1579 *** (0.0125)	-0.1439 *** (0.0125)	-0.1589 *** (0.0128)	-0.1652 *** (0.0122)	-0.1591 *** (0.0115)
サンプル数	1,318.81	1,047.38	1,426.57	1,439.64	1,406.56	1,638.12	1,659.28	1,462.18	1,153.13	1,560.87	1,343.60	1,768.92
Wald χ^2 square	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Prob > χ^2 square	0.0103	0.0069	0.0098	0.0107	0.0106	0.0111	0.0107	0.0119	0.0101	0.0111	0.0107	0.0104
擬似決定係数												

2019年

被説明変数評価ポイント

相計方法 Orderedprobit regression

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
使い道 (ブレゼント: 1, それ以外: 0)	0.4121 *** (0.0596)	0.3676 *** (0.0662)	0.3206 *** (0.0657)	0.2975 *** (0.0702)	0.3918 *** (0.0557)	0.2504 *** (0.0537)	0.4792 *** (0.0566)	0.4234 *** (0.0675)	0.3313 *** (0.0693)	0.6062 *** (0.0820)	0.3652 *** (0.0628)	0.2925 *** (0.0516)
頻度 (リピーター: 1, それ以外: 0)	0.2471 *** (0.0169)	0.2417 *** (0.0169)	0.3195 *** (0.0164)	0.3369 *** (0.0165)	0.3565 *** (0.0165)	0.3629 *** (0.0149)	0.3593 *** (0.0155)	0.3593 *** (0.0181)	0.3800 *** (0.0167)	0.3593 *** (0.0174)	0.3529 *** (0.0162)	0.3520 *** (0.0148)
語句1 ("エレガン", "ゴージャズ", "華やか", "豪華", が含まれるケース: 1, それ以外: 0)	0.4989 *** (0.0515)	0.4425 *** (0.0561)	0.4967 *** (0.0485)	0.4847 *** (0.0548)	0.5823 *** (0.0653)	0.5728 *** (0.0590)	0.4670 *** (0.0583)	0.4909 *** (0.0678)	0.5042 *** (0.0551)	0.5823 *** (0.0584)	0.5823 *** (0.0546)	0.5085 *** (0.0511)
語句2 ("可愛い", "かわいい") が含まれるケース: 1, それ以外: 0)	0.2345 * (0.1277)	0.4137 *** (0.1610)	0.4451 *** (0.1321)	0.2906 (0.1458)	0.3791 *** (0.1401)	0.1899 (0.1095)	0.3978 *** (0.1218)	0.1889 (0.1255)	0.3096 ** (0.1245)	0.3457 *** (0.1263)	0.1744 (0.1331)	0.0130 (0.1434)
語句3 ("可愛い", "かわいい") が含まれるケース: 1, それ以外: 0)	0.2298 *** (0.0144)	0.1801 *** (0.0162)	0.1600 *** (0.0140)	0.2217 *** (0.0147)	0.2112 *** (0.0149)	0.2131 *** (0.0132)	0.2106 *** (0.0133)	0.2295 *** (0.0150)	0.1378 *** (0.0142)	0.2016 *** (0.0155)	0.2094 *** (0.0144)	0.2047 *** (0.0134)
語句4 ("キレイ", "きれい", "美しい", "美容", "綺麗", が含まれるケース: 1, それ以外: 0)	0.3754 *** (0.0184)	0.3213 *** (0.0200)	0.3379 *** (0.0170)	0.3458 *** (0.0183)	0.3099 *** (0.0191)	0.3456 *** (0.0175)	0.3212 *** (0.0182)	0.2819 *** (0.0207)	0.3430 *** (0.0190)	0.3471 *** (0.0188)	0.3583 *** (0.0185)	0.3478 *** (0.0171)
語句5 ("カッコイイ", "カッコーイ") が含まれるケース: 1, それ以外: 0)	0.6536 ** (0.2003)	0.4925 ** (0.1932)	0.6094 *** (0.2054)	0.6990 *** (0.1944)	0.3374 (0.2111)	0.5650 ** (0.2191)	0.2715 (0.2255)	0.4025 * (0.2216)	0.7817 *** (0.2323)	0.3117 (0.2055)	0.8624 *** (0.2054)	0.5758 *** (0.2121)
語句6 ("オシャレ", "おしゃれ", "お洒落") が含まれるケース: 1, それ以外: 0)	0.3902 *** (0.0445)	0.2978 *** (0.0460)	0.3683 *** (0.0414)	0.4454 *** (0.0446)	0.4099 *** (0.0443)	0.4248 *** (0.0397)	0.3461 *** (0.0399)	0.3756 *** (0.0449)	0.3319 *** (0.0404)	0.3504 *** (0.0434)	0.4137 *** (0.0401)	0.4140 *** (0.0382)
語句7 ("センス", "センス", "センス") が含まれるケース: 1, それ以外: 0)	0.0088 (0.0270)	-0.0875 (0.0285)	0.0034 (0.0260)	-0.0191 (0.0274)	0.0171 (0.0237)	-0.0570 ** (0.0241)	0.0067 (0.0283)	-0.0288 (0.0283)	-0.0183 (0.0283)	-0.0075 (0.0282)	-0.0222 (0.0254)	-0.0175 (0.0236)
語句8 ("デザイン", "デザイン", "デザイン") が含まれるケース: 1, それ以外: 0)	-0.0088 *** (0.0161)	-0.0775 *** (0.0177)	-0.1857 *** (0.0153)	-0.1159 *** (0.0165)	-0.1392 *** (0.0166)	-0.1294 *** (0.0148)	-0.1007 *** (0.0152)	-0.0924 *** (0.0168)	-0.1042 *** (0.0162)	-0.1298 *** (0.0169)	-0.1393 *** (0.0162)	-0.1010 *** (0.0153)
語句9 ("ベージュ", "ベージュ", "ベージュ") が含まれるケース: 1, それ以外: 0)	-0.0151 (0.0271)	0.0032 (0.0296)	-0.0485 ** (0.0240)	-0.0418 (0.0257)	-0.0479 * (0.0281)	-0.0989 *** (0.0242)	-0.1021 *** (0.0259)	-0.1037 *** (0.0265)	-0.0405 (0.0276)	-0.0660 (0.0270)	-0.0172 (0.0255)	0.0351 (0.0236)
語句10 ("赤", "レッド") が含まれるケース: 1, それ以外: 0)	-0.0231 (0.0393)	-0.0298 (0.0451)	0.0333 (0.0442)	-0.0423 (0.0461)	-0.0203 (0.0502)	0.0228 (0.0401)	0.0264 (0.0438)	0.0986 ** (0.0470)	-0.0989 ** (0.0429)	-0.0828 * (0.0462)	0.0372 (0.0438)	0.0582 (0.0665)
語句11 ("ピンク", "ピンク", "ピンク") が含まれるケース: 1, それ以外: 0)	0.0088 (0.0345)	-0.1234 *** (0.0381)	-0.1463 *** (0.0391)	-0.0826 (0.0323)	-0.0282 (0.0340)	-0.0139 (0.0316)	-0.0841 *** (0.0324)	-0.0701 * (0.0376)	-0.1062 *** (0.0362)	-0.1080 *** (0.0389)	-0.0188 (0.0375)	0.0168 (0.0348)
語句12 ("イエロー", "黄色") が含まれるケース: 1, それ以外: 0)	-0.1237 ** (0.0576)	-0.0551 (0.0667)	-0.3654 *** (0.0532)	-0.1930 *** (0.0510)	-0.0827 (0.0507)	-0.1545 *** (0.0417)	-0.0669 (0.0464)	0.0231 (0.0517)	-0.1503 ** (0.0583)	-0.0639 (0.0652)	-0.2032 *** (0.0666)	-0.1219 ** (0.0613)
語句13 ("グリーン", "緑") が含まれるケース: 1, それ以外: 0)	0.0965 *** (0.0360)	0.0987 ** (0.0443)	-0.0231 (0.0356)	0.0396 (0.0390)	0.0436 (0.0388)	0.0176 (0.0338)	0.0139 (0.0342)	0.0231 (0.0409)	0.0519 (0.0402)	0.0432 (0.0397)	0.0366 (0.0393)	0.0548 (0.0355)
語句14 ("ブルー", "ブルー") が含まれるケース: 1, それ以外: 0)	0.0573 *** (0.0224)	0.0398 *** (0.0262)	-0.0774 *** (0.0225)	0.0859 *** (0.0249)	0.0503 * (0.0269)	0.0329 (0.0247)	0.0496 * (0.0281)	0.0158 (0.0317)	0.0541 ** (0.0275)	0.0891 *** (0.0266)	0.0781 *** (0.0232)	0.0666 *** (0.0212)
語句15 ("黒", "ブラック") が含まれるケース: 1, それ以外: 0)	0.0585 *** (0.0171)	0.0504 *** (0.0189)	0.157 *** (0.0155)	0.0838 *** (0.0185)	0.0987 *** (0.0185)	0.1003 *** (0.0165)	0.1112 *** (0.0168)	0.1250 *** (0.0190)	0.0927 *** (0.0183)	0.1058 *** (0.0190)	0.0832 *** (0.0173)	0.0690 *** (0.0155)
語句16 ("肌ざわり", "手触り") が含まれるケース: 1, それ以外: 0)	0.1374 ** (0.0584)	0.1883 *** (0.0632)	0.1750 *** (0.0652)	0.1610 ** (0.0716)	0.2323 *** (0.0757)	0.1338 ** (0.0664)	0.1883 *** (0.0707)	0.2394 *** (0.0805)	0.1773 *** (0.0638)	0.1749 *** (0.0592)	0.2367 *** (0.0527)	0.1488 *** (0.0465)
語句17 ("生地") が含まれるケース: 1, それ以外: 0)	-0.1303 *** (0.0125)	-0.1828 *** (0.0137)	-0.1455 *** (0.0117)	-0.1329 *** (0.0120)	-0.1294 *** (0.0120)	-0.1302 *** (0.0106)	-0.1308 *** (0.0115)	-0.1223 *** (0.0126)	-0.1223 *** (0.0120)	-0.0845 *** (0.0126)	-0.1380 *** (0.0120)	-0.1283 *** (0.0112)
サンプル数	50,686	41,573	53,407	47,203	47,316	58,311	55,061	45,153	48,763	43,326	50,324	60,553
Wald χ^2 square	1,472.30	975.14	1,452.69	1,448.03	1,366.78	1,732.45	1,525.20	1,115.98	1,330.03	1,271.04	1,544.73	1,672.60
Prob > χ^2 square	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
擬似決定係数	0.0106	0.0688	0.0100	0.0103	0.0108	0.0111	0.0103	0.0098	0.0104	0.0109	0.0111	0.0103

5. 機械学習（教師あり）を用いた評価ポイントに関する分析の評価

前章では評価ポイントに関する統計分析を行った。本章では教師ありの機械学習を用いて、前章で用いた説明変数によってどの程度、被説明変数である評価ポイントを見分けることができるかどうかを検証した。データの70%を学習用に用いて作成したモデルを用いて、残り30%の評価ポイントを予測してその正解率を計算した。

Python のライブラリーのなかの scikit-learn を用いて、SVM (Support Vector Machine)、決定木、ランダムフォレスト、k 近傍法といった教師あり機械学習を行った。SVM は誤判別をなるべく防ぐために境界とデータの距離であるマージンを最大化するような境界を引いて分類するモデルである。決定木は段階的にデータを分割していき、木が枝分かれするように分析していく。決定木はデータを分割していく流れを把握することができるため解釈が容易である。ランダムフォレストは決定木を複数並列に作成して、各決定木の出力結果の多数決によって予測を行う。k 近傍法は、与えられた学習データをベクトル空間上にプロットしておき、未知のデータが得られたら、そこから距離が近い順に任意の k 個を取得し、その多数決でデータが属するクラスを推定する。

それぞれのモデルには特徴があるので本稿では、まず評価ポイント（1～5 の 5 段階）について SVM、決定木、ランダムフォレスト、k 近傍法を用いた。それぞれ学習データを 70% として学習によってモデル構築をして、次の段階で残り 30% について構築したモデルによる予測を行い、その正答率を計算した。その結果が図表 4 である。年月による違いはあまり見られず正答率は 50% 未満となった。SVM、決定木、ランダムフォレストの結果はほぼ同じであり、k 近傍法が最も低い正答率となった。次に評価ポイントのなかで 4 と 5 を併せて 1 つにし、1～3 も併せて 1 つにして 2 値モデルとした。同じ方法を用いて正答率を計算した（図表 5）。SVM、決定木、ランダムフォレストについては正答率が 70% 台まで上昇した。k 近傍法もそれ程でないにしても評価ポイントが 5 段階の図表 4 と比較すると正答率が上昇していることがわかる。

正答率を上げるためには、説明変数に工夫が必要であると考え。特にクチコミをもとにした変数については該当しているか否かの 1.0 の値となっている。TF-IDF を用いることや、特定の語句とのコサイン類似度を用いることなどは今後の課題としたい。それ以外に、クチコミのなかでまだ採用していない語句を分析に取り入れることも必要だと考える。

図表4 教師あり機械学習による正答率(評価ポイント5段階)

2015年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.443487	0.455003	0.450240	0.451518	0.450597	0.447351	0.452664	0.450703	0.444688	0.440728	0.436096	0.441060
決定木	0.439798	0.452586	0.449488	0.446498	0.447800	0.445934	0.449389	0.447943	0.441285	0.436789	0.457870	0.438829
ランダムフォレスト	0.441035	0.453416	0.450472	0.448029	0.448687	0.447223	0.450741	0.448972	0.442570	0.437861	0.436546	0.441138
k近傍法	0.402868	0.299592	0.378306	0.386434	0.247959	0.376613	0.273563	0.418387	0.349863	0.276035	0.229402	0.362841

2016年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.450740	0.471901	0.482075	0.457176	0.458333	0.465824	0.468436	0.461838	0.463018	0.460231	0.451063	0.448245
決定木	0.444761	0.463902	0.473527	0.450187	0.451102	0.459768	0.462017	0.455142	0.451851	0.451877	0.442633	0.439919
ランダムフォレスト	0.447137	0.466245	0.475188	0.452146	0.453592	0.461658	0.464972	0.458080	0.455885	0.454191	0.445056	0.442615
k近傍法	0.266017	0.357161	0.241624	0.374555	0.403153	0.412004	0.423855	0.427537	0.239841	0.337172	0.400989	0.409216

2017年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.458259	0.482646	0.474161	0.469491	0.475043	0.460547	0.461507	0.469328	0.471544	0.462871	0.454314	0.441859
決定木	0.451718	0.475132	0.467106	0.459621	0.466781	0.450760	0.450182	0.461691	0.457887	0.453473	0.445193	0.434550
ランダムフォレスト	0.455046	0.478618	0.469457	0.462791	0.469604	0.453070	0.453644	0.463681	0.461652	0.455562	0.448366	0.436660
k近傍法	0.347696	0.432301	0.432301	0.340105	0.354286	0.354286	0.213883	0.403298	0.404370	0.338001	0.299043	0.373655

2018年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.452412	0.471491	0.475753	0.465942	0.463913	0.462466	0.469533	0.471087	0.462922	0.444870	0.446722	0.444426
決定木	0.444673	0.463367	0.465176	0.456324	0.457674	0.454686	0.464089	0.461605	0.457230	0.438062	0.438122	0.438098
ランダムフォレスト	0.446358	0.466000	0.468001	0.459223	0.460563	0.457537	0.466256	0.463903	0.458672	0.440045	0.440709	0.441345
k近傍法	0.388379	0.442606	0.353404	0.272464	0.405136	0.310963	0.307282	0.351914	0.399772	0.359268	0.325100	0.403048

2019年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.443487	0.477224	0.470348	0.465514	0.475712	0.474427	0.468210	0.485522	0.472587	0.465331	0.460822	0.454218
決定木	0.440357	0.468320	0.464906	0.458231	0.466778	0.467686	0.460353	0.480550	0.465548	0.465730	0.451184	0.447779
ランダムフォレスト	0.442071	0.471580	0.463298	0.461032	0.469779	0.469102	0.462752	0.482890	0.468323	0.459006	0.454331	0.449798
k近傍法	0.404762	0.225614	0.414322	0.343533	0.330123	0.402775	0.409429	0.410573	0.433261	0.387610	0.410137	0.343610

図表5 教師あり機械学習による正答率(評価ポイント2段階)

2015年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.768844	0.791379	0.791566	0.784957	0.778829	0.782538	0.794490	0.788879	0.777561	0.769252	0.769935	0.760339
決定木	0.767389	0.789646	0.789849	0.782728	0.776578	0.780697	0.793080	0.787279	0.776228	0.767155	0.768204	0.758309
ランダムフォレスト	0.768366	0.790402	0.791123	0.784169	0.777601	0.781710	0.793899	0.787893	0.777085	0.768273	0.769058	0.759502
k近傍法	0.724834	0.550148	0.789618	0.775231	0.776464	0.768140	0.538711	0.782130	0.728971	0.720679	0.543354	0.752905

2016年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.756553	0.780354	0.782700	0.763221	0.761913	0.764327	0.765527	0.765015	0.757703	0.753034	0.739741	0.735642
決定木	0.753188	0.776838	0.779379	0.760061	0.757112	0.760742	0.764327	0.759002	0.752441	0.748688	0.736056	0.730864
ランダムフォレスト	0.754574	0.778876	0.781039	0.761574	0.758950	0.762728	0.763285	0.762692	0.755365	0.750889	0.737772	0.733513
k近傍法	0.544587	0.773832	0.746557	0.666934	0.737316	0.742382	0.737861	0.737889	0.746653	0.738357	0.713089	0.714306

2017年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.746801	0.771150	0.766653	0.752684	0.757108	0.751368	0.754313	0.753337	0.762752	0.746699	0.743330	0.727081
決定木	0.742957	0.766346	0.762008	0.747929	0.753184	0.746322	0.751144	0.749615	0.757511	0.740986	0.738911	0.727081
ランダムフォレスト	0.745137	0.768826	0.764536	0.750666	0.754699	0.748389	0.753315	0.751348	0.759873	0.743259	0.741120	0.725279
k近傍法	0.720638	0.764565	0.756247	0.748073	0.718417	0.738419	0.526171	0.698601	0.756404	0.743259	0.541664	0.653369

2018年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.737065	0.773131	0.764437	0.753623	0.742103	0.746526	0.753874	0.761224	0.752562	0.730235	0.734514	0.720480
決定木	0.733134	0.769219	0.761012	0.748946	0.739279	0.743259	0.750319	0.757058	0.747628	0.725013	0.729794	0.716463
ランダムフォレスト	0.734382	0.771325	0.762514	0.751647	0.740724	0.744685	0.751597	0.759213	0.750209	0.727129	0.732122	0.718444
k近傍法	0.688136	0.765157	0.747912	0.557839	0.727852	0.687790	0.727768	0.737663	0.735636	0.690970	0.722553	0.703588

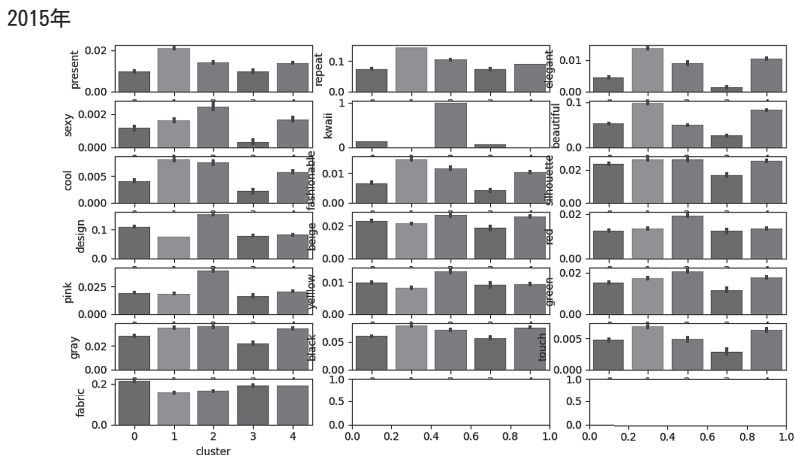
2019年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
SVC	0.737103	0.770093	0.760868	0.752188	0.755863	0.753101	0.749820	0.770328	0.765060	0.751143	0.751143	0.737204
決定木	0.733650	0.765880	0.757405	0.749457	0.752303	0.749080	0.746041	0.767768	0.760187	0.746647	0.746647	0.734530
ランダムフォレスト	0.735409	0.767708	0.759941	0.750788	0.754118	0.750722	0.748560	0.768865	0.762217	0.749238	0.749238	0.736331
k近傍法	0.729742	0.556722	0.737926	0.726420	0.744207	0.735882	0.733745	0.745320	0.754501	0.695748	0.695748	0.720452

6. 非階層クラスター分析

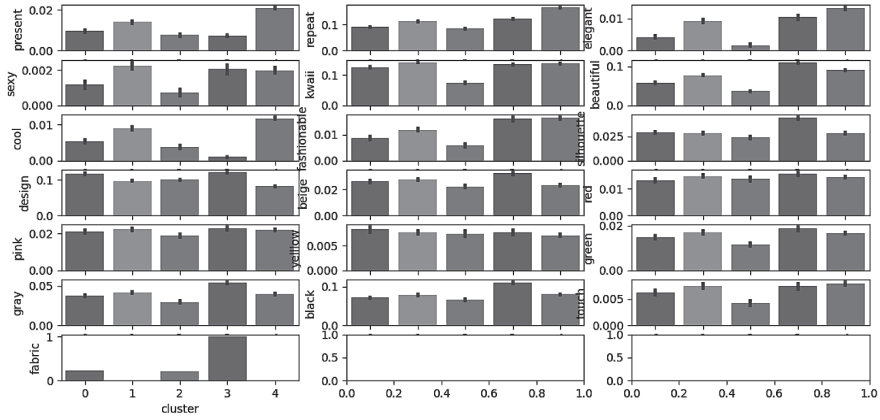
本章では4章、5章に引き続き、評価ポイントに関する分析について考察を進める。評価ポイントを推計する際に用いた説明変数について、教師なしの機械学習のなかの非階層クラスター分析を行う。教師ありの際と同様、Pythonのライブラリーのなかのscikit-learnを用いる。非階層クラスター分析では事前にクラスター数を指定する必要がある、クラスター数を3～5に設定した結果、最もクラスター毎の特徴が出たのが、クラスター数5の場合であったのでその結果を紹介する（図表6）。2015年から2019年の年単位の分析である。

グラフの横軸はcluster番号（0～4）、縦軸が各変数の寄与を表している。2015年からみていくとcluster0については、black（黒）の寄与が0.2程度、design（デザイン）の寄与は0.1となっている。Cluster1はblack（黒）の寄与が0.2弱、repeat（リピーター）の寄与が0.1超となっている。Cluster2については、kawaii（かわいい）の寄与が1程度と圧倒的な結果となった。Cluster3についてはblack（黒）の寄与が0.2弱となっている。Cluster4についてもblack（黒）の寄与が0.2弱と最も高くなっている。以上でわかるように、black（黒）、design（デザイン）、repeat（リピーター）、kawaii（かわいい）が他を圧倒している。2016年に特徴的であったのは、black（黒）の寄与がCluster3に集中していることであり、2017年と2018年も同様である。2019年はkawaii（かわいい）の寄与がCluster0に集中していることである。この傾向は2015年と同様である。

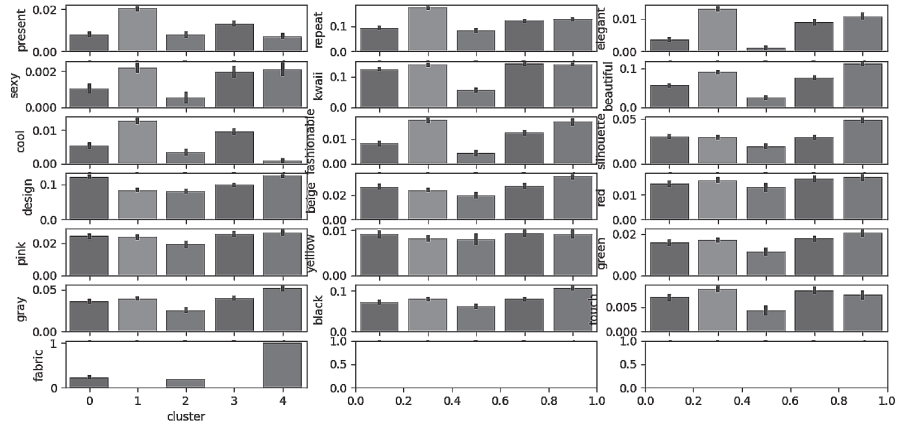
図表6 非階層クラスター分析の結果



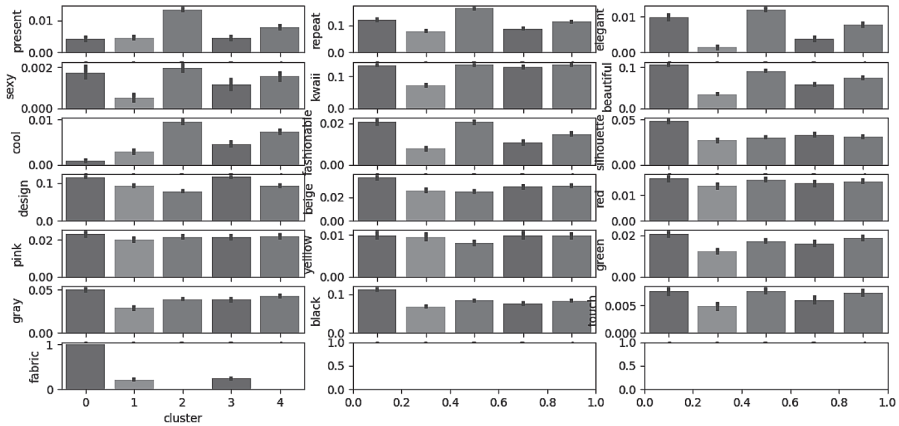
2016年



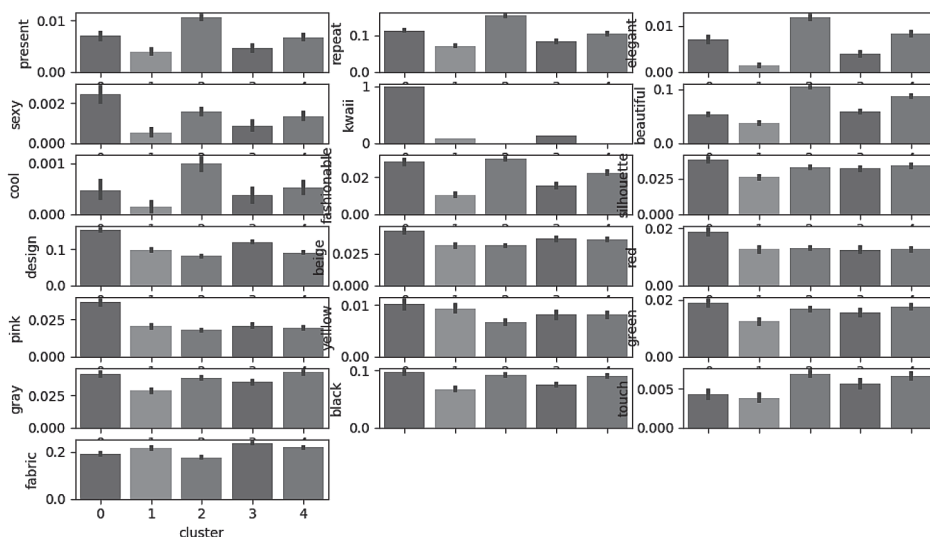
2017年



2018年



2019年



7. お わ り に

本稿ではレディースファッションのクチコミの分析を形態素解析の結果の月次単位のグラフ化、「おしゃれ」に関するコサイン類似度の計算，評価ポイントに関する統計分析と教師なしの機械学習のなかの非階層クラスター分析を行った。

形態素解析の結果の月次単位のグラフ化の結果、「色」は6月を挟んだ2カ月，つまり4月から8月に値が大きい傾向にあり，「美」は3月，4月に多くみられ，「華やか・豪華」は1月から3月に相対的に多い傾向にあることがわかった。「かわいらしさ」は7月，8月と多くなっており，「視覚に関する用語」は，7月前後が多い傾向にあり，「季節感」は傾向として3月，4月と夏に多くみられる。「場・機会」は3月，4月と，7月から9月と年2回のピークがあることがわかった。レディースファッションに関するクチコミには，こうした季節的な特徴があることが明らかとなった。

本稿では，「おしゃれ」という語句と他の語句のコサイン類似度を年ごとに計算した。「おしゃれ」という語句と類似度が相対的に高い語句を探索した。「似合う」という語句はどの年次もプラスの値であり，比較的安定して「おしゃれ」と類似度が高かった。年別にみると，2015年はフランネル，似合う，ネイビー，カッコイイ，ステキが高い類似度になった。2016年はステキ，コーディネート，フランネル，グレー，ラベンダー，合うといった語句が相対的に高い値となった。2017年はかっこいい，モスグリーン，合う，似合

う、カッコイイが相対的に高い値となった。2018年は化粧、写真、時期、かっこいい、びっくりが相対的に高い値となった。2019年は手ざわり、大事、感謝、サプライズ、信頼が高い値となった。2019年は感情や気持ちに関する語句が「おしゃれ」と高い類似度となっていることがわかった。

クチコミ投稿者がつけた評価ポイントについて順序プロビットモデルを用いて推計した結果、1%水準で統計的に有意であった説明変数は、使い道（プレゼント）、頻度、語句1（エレガント、ゴージャス、華やか、豪華）、語句3（可愛い、かわいい）、語句6（オシャレ、おしゃれ、お洒落）、語句8（デザイン）、語句17（生地）である。符号がマイナスであった変数は、語句8（デザイン）、語句17（生地）である。マイナスになった理由として考えられるのは、デザインや生地に関して不満があったクチコミであったのではないかと推察される。係数の大きさについては、語句1（エレガント、ゴージャス、華やか、豪華）、使い道、頻度といった変数が他と比較して大きく、評価ポイントに影響をもっていたことがわかった。

評価ポイントについて、SVM (Support Vector Machine)、決定木、ランダムフォレスト、k近傍法といった教師あり機械学習を行った。正答率を計算した結果、年月による違いはあまり見られず正答率は50%未満となった。SVM、決定木、ランダムフォレストの結果はほぼ同じであり、k近傍法が最も低い正答率となった。次に評価ポイントのなかで4と5を併せて1つにし、1～3も併せて1つにして2値モデルとして正答率を計算した結果、SVM、決定木、ランダムフォレストについては正答率が70%台まで上昇した。

評価ポイントについて教師なしの機械学習のなかの非階層クラスター分析の結果、年によって相違はあるもののblack（黒）、design（デザイン）、repeat（リピーター）、kawaii（かわいい）がクラスターの形成に寄与していることがわかった。

全体を通して様々な分析手法を組み合わせることの必要性を改めて認識した。レディースファッションに関するクチコミの性格上、季節性や流行があり、月単位、年単位の分析の必要性もあることがわかった。本稿の一つの大きな目的として、「おしゃれ」という抽象的な語句がどのような語句と近い意味を持つかを検証することであった。メンズファッションやアクセサリーといった、他の分野のクチコミを同様に分析して得られた結果と本稿の結果を比較したい。こうした抽象的な語句の分析は他の分野のクチコミにも用いることができると考える。本稿のもう一つの主たる目的は、評価ポイントに関する分析であった。従来型の統計分析に加えて、機械学習を取り入れたが、変数選択やクラスター数の決定に分析者の視点を多く取り入れた。今後、機械学習や深層学習によってカテゴリー分類

の自動化にも取り組みたい。

クチコミの分析は機械学習や深層学習によって、コンピューターによる自動化が進んでいる。他方、手法の選択は分析者が行う必要性もあるのではないかと考える。今後の研究としては本稿で残された課題に取り組むとともに、レディースファッション以外にも、様々なクチコミの分析を進めていき、分野による相違点や共通点に焦点をあてたい。

謝 辞

本研究では、国立情報学研究所の IDR データセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」(https://rit.rakuten.com/data_release/) のなかの「楽天市場」を利用した。感謝したい。

参 考 文 献

- 会川雄大・森本千佳子 (2021), 「宿泊サイトの口コミにおける数値評価の信頼性検証」, 経営情報学会 全国研究発表大会要旨集 202111(0), pp.123-124, 2022-01-31
- 石橋太郎 (2012), 「e- 口コミのテキスト・マイニング分析に向けて (その 1) : 伊豆地域におけるホテル・旅館を対象として」, 静岡大学経済研究, 17巻 2号 pp.1-11
- 大川慎 (2019), 「テキストマイニングを利用した観光地イメージの言語間比較に関する試み」, 2019年39巻 2号 pp. 85-94
- 久保田貴文, 「旅行情報サイトにおけるホテル口コミデータのテキストマイニング」, 経営・情報研究 多摩大学研究紀要, 2016.
- 酒井美春・松下光範・上田真由美, 「化粧品の評価項目別スコア生成のための評価表現辞書の自動構築」, DEIM Forum 2019 B6-2.
- 白田由香利, 橋下隆子・久保山哲二 (2013), 「インターネット上の口コミサイトにおける化粧品の評判分析」, 学習院大学計算機センター年報, No.33, pp.2-7, 2013.
- 杉本徹雄編 (1997), 『消費者理解のための心理学』福村出版, 1997 p.1-237.
- 杉谷陽子 (2009), 「インターネット上の口コミの有効性: 製品の評価における非言語的の手がかりの効果」, 『上智経済論集』54(1・2), pp. 47-58.
- 高野敦子・池奥渉太・北村泰彦 (), 「因果関係に着目した口コミ Web サイトからの評価表現抽出」, 人工知能学会論文誌2009年24巻 3号 pp. 322-332.
- 高島侑里, 青野雅樹 (2017), 「化粧品レビューサイトにおけるクチコミの有用性判定」, 言語処理学会 第23回年次大会発表論文集, pp.799-802, 2017.
- 立花伸一・原田智彦・津田和彦 (2021), 「テキストマイニングを用いたレシピ口コミ分析による味覚表現の意図解明」, 2021年141巻10号, pp. 1115-1124.
- 辻井康一・津田和彦 (2012), 「テキストマイニングを用いた宿泊レビューからの注目情報抽出方法」, 情報処理学会デジタルプラクティス Vol. 3 No. 4 October 2012, pp.289-296.
- 中村裕太・佐藤弘喜 (2018), 「料理レシピサイト内のクチコミのテキスト解析 クチコミの信頼性に関する研究」, 日本デザイン学会 第65回春季研究発表大会, pp.16-17.
- 野守耕爾・神津友武 (2015), 「口コミデータにPLSA を適用した観光客目線による観光地分析」人工知能学会全国大会論文集, pp.1-4.
- 濱岡豊 (1994), 「クチコミの発生と影響のメカニズム」, 消費者行動研究 Vol. 2 No. 1 (1994, 9), pp.29-73.
- 松波友稀・上田真由美・中島伸介 (2017), 「コスメアイテムに対する評価項目別レビュー自動スコア

- リング方式の開発」, 第9回データ工学と情報マネジメントに関するフォーラム, B5-3, 2017.
- 松本義之 (2016), 「SNS から収集した近接地域情報のテキストマイニングによる分析」, バイオメディカル・ファジィ・システム学会誌 Vol.18, No. 2, pp.41-48 (2016).
- Arndt, Johan (1967), “Role of Product Related Conversations in the Diffusion of a New Product”, *Journal of Marketing Research*, Vol. 4, pp.291-29.