

評価付きレビューにおけるトピックモデルの提案

金 勝 鎮

2024年2月29日

要旨

本研究は、オンラインレビューを分析対象とし、自然言語で記述されたレビュー文書と整数値で付されたレビュー得点との関係を同時に分類するトピックモデルを提案した。提案モデルは自然言語の分類モデルである Latent Dirichlet Allocation (LDA) を拡張するモデルであり、整数値であるレビューに付されている満足スコアを分類するためにディリクレ-二項分布を仮定した Binomial Factor Model (BFM) を開発し、LDA と BFM を合成させて同時分類を行っている。モデルは全て共役な分布を用いて定義されているため、collapsed gibbs sampling でトピックを推定することができ、効率的かつ安定した結果を得ることができる。また、文書の語と整数値のスコアを同時分類することができるため、所属されるトピックから特定の語の観測確率とレビューの満足スコアとの関係を考察することができる。

キーワード

トピックモデル、オンラインレビュー、CGM、テキスト分析

1 研究背景

消費者生成メディア (CGM; consumer generated media) による情報共有は、企業のマーケティング戦略における重要な情報源として認識されている。Kannan and Li (2017) によると、デジタルマーケティングの顕著な特徴の一つは、消費者の購買プロセスを詳細に追跡できることである。しかしながら、購買履歴データからは購買前後の消費者の心理的側面を明らかにすることが困難である。そこで、CGM が消費者行動研究においても有用であることが、Lee and Bradlow (2011) および Humphreys and Wang (2018) によって示唆されている。

オンラインレビューは、CGMの中でも特に、企業と消費者双方にとって重要な情報源として機能している。製品の実際の売上や企業の業績との関連性については、Chevalier and Mayzlin (2006) はオンライン上の書籍について書かれたレビューがオンライン書籍の売上に与える影響を示し、Liu (2006)、Onishi and Manchanda (2012) はオンラインレビューが映画の売上に影響を及ぼしていることを明らかにした。また、企業の総合的な評価、例えば株価との関連性についても、Luo (2009)、Tirunillai and Tellis (2012) によって研究されている。

オンラインレビューの量的な側面だけでなく、質的な側面にも注目が必要である。Stephen and Galak (2012) は、支払いを伴うメディア (paid media; ペイドメディア)、企業が所有するメディア (owned media; オウンドメディア)、そして他者や消費者が生成するメディア (earned media; アーンドメディア) という三つのカテゴリーにメディアを分類している。ペイドメディアやオウンドメディアは企業のコントロール下にあるが、消費者が生成するメディアであるアーンドメディア、いわゆる CGM は企業が直接コントロールできないメディアであるため、ネガティブなレビューが発信されるリスクを含んでいる。Anderson and Simester (2014) は、ネガティブなレビューが消費者の製品に対する認識や購買行動にどのように影響を与えるかを示している。彼らの研究では、ネガティブなフィードバックが実際には製品の改善につながり、長期的にはポジティブな効果をもたらす可能性があることを示唆した。一方、Chen and Lurie (2013) は、オンラインレビューの情報量が消費者の製品評価や購買意欲にどのように影響を及ぼすかに焦点を当てている。彼らは、レビューの情報量が多いほど、消費者が製品に対してより信頼を寄せ、購買意欲が高まることを明らかにした。

したがって、CGM の分析にあたっては、情報の量だけでなく、情報の感情的傾向 (sentiment や valence) を考慮することが重要である。感情の傾向に関しては、辞書を用いた分析が一般的であるが、製品カテゴリーによって感情を示す単語が異なる可能性があるため、一般的な辞書を用いた分析には限界があるとされる。

本研究では、レビューに付された得点 (score) とレビュー本文に現れる単語との関連性を分析することで、マーケティングリサーチや製品開発に寄与する情報を提供することを目指す。そのため、レビュー本文のような自然言語を定量化し、レビュースコアとの関係を分析するためのモデルの開発が求められている。この分野における先

行研究としては、Pang and Lee (2008) が感情分析のための自然言語処理技術の応用に関する研究などが行っており、本研究はこれらの研究を基盤として、オンラインレビューの得点と本文の両方を考慮した分類モデルの提案を行う。

2 先行研究

2.1 自然言語分析：Natural Language Analysis

消費者によって生成された自然言語のデータを活用する試みは、マーケティング研究の分野で高い関心を集めており、多数の研究が行われている (Berger et al., 2019; Humphreys and Wang, 2018; Lamberton and Stephen, 2016)。先に述べた通り、オンラインレビューの分析に際しては、量的な分析だけでは不十分であり、質的な検討も同時に行う必要がある。これは、単純な数値集計を超えて、テキスト内容そのものの理解を要求されるためである。

CGM における自然言語データは、その性質上、非構造化データに分類される (Balducci and Marinova, 2018)。この非構造化データを適切に定量化するための方法論については、Berger et al. (2019) が詳細な分析手法を提供している。彼らは、テキストデータを分析するための手法として、3つのアプローチを紹介している。第一に、単語そのものから質的な側面を分析する手法があり、LIWC (Linguistic Inquiry and Word Count) を用いた感情分析がその一例である (Berger and Milkman, 2012; Hewett et al., 2016)。この手法を用いることで、テキストからネガティブな感情やポジティブな感情を持つレビューを区別することが可能となる。第二に、テキストからトピックを抽出する手法がある。このアプローチでは、外部からの基準を設けずにトピックを探索的に分類し、テキストデータの解釈や新たな知見を得ることができる。第三に、テキスト内の単語間の関係性を分析する手法が挙げられる。本研究では、特にトピック抽出に焦点を当て、そのためのモデル構築と分析を行う。この目的のために、次節ではトピックモデルの一つである LDA (Latent Dirichlet Allocation) に関する研究を概観する。

2.2 LDA の拡張

Latent Dirichlet Allocation (LDA) とその拡張モデルは、マーケティング研究

におけるテキスト分類の手法として広く用いられている。LDA モデルは Blei and Jordan (2003) によって開発され、文書内の単語が高次元のカテゴリカル分布に従うと仮定し、その共役事前分布としてディリクレ分布を採用している。機械学習の分野は、このモデルの拡張が積極的に行われており、著者の影響を考慮する Author Topic Model (Rosen-Zvi et al., 2004) や、時系列データを扱う Dynamic Topic Model (Blei and Lafferty, 2006)、さらには階層的なトピック構造を可能にする Pachinko Allocation Model など、多様なバリエーションが提案されている。また、LDA モデルは、Griffiths and Steyvers (2004) によって提案された collapsed gibbs sampling による推定が可能であり、これは Markov-Chain Monte Carlo (MCMC) 法を用いた効率的な推定手法である。

マーケティングの領域では、Liu et al. (2017) が Twitter データの分析に LDA を適用し、Tirunillai and Tellis (2014) はオンラインレビュー分析における LDA の応用を示している。さらに、Büschken and Allenby (2016) は Author Topic Model を拡張したモデルの分析結果を報告し、Nam et al. (2017) でも LDA の応用が見られる。LDA は多項分布を前提とする分析対象であれば、自然言語データに限らず適用可能であり、Trusov et al. (2016) はウェブサイト訪問行動の研究に Correlated Topic Model (CTM) を拡張したモデルを用いている。また、Toubia et al. (2019) による研究では、CGM データへの LDA の適用が示されている。

LDA の推定において collapsed gibbs sampling を利用することで、条件付き事後分布からのサンプリングすることが可能となり、モデルの拡張が容易になる。一部の研究では、パラメータ推定に Metropolis-Hastings (M-H) 法を用いているが (Trusov et al., 2016; Büschken and Allenby, 2016)、M-H 法は計算負荷が高く、推定結果の安定性を確保することが難しい場合がある。そこで、本研究では、レビュースコアが連続的な整数値を取り、最大値が存在するという特性を活かし、二項分布を事前分布とする新たなモデルを提案する。

3 モデル

3.1 表記の定義

本節ではモデルの概要を示す。モデルの説明では、まず LDA モデルを検討し、加

えて二項分布を仮定した Binomial Factor Model を定義する。その上で提案モデルである Score-Text Allocation Model を提示する。

まず、レビューデータの表記について定義する。レビューの件数は D 個存在し、各レビューは *document* と呼ばれる。レビュー d には自然言語で書かれた文章と得点が付されている。文章について、データセットから得られるすべての語数は N とおく。ここで、第 i 番目 ($i = 1, \dots, N$) で観測される語を w_i とおく。観測される語の種類を V とおき、これを *vocabulary* とよぶと、 i 番目の語としては V 種類ある単語のいずれか1つが観測されることになる。したがって、 w_i は V 次元のベクトルとなり、ここで語彙 v が観測されたときは $w_{iv} = 1$ となり、 $v' \neq v$ のとき $w_{iv'} = 0$ となる変数として定義できる。また、全ての語は D 個あるいずれか1つの文書に所属することになる。この所属関係を示す変数として x_i を導入する x_i は D 次元のベクトルであり、第 i 語が文書 d に所属していたときには $x_{id} = 1$ をとり、 $d' \neq d$ について $x_{id'} = 0$ をとる変数として定義できる。

次に得点について定義する。レビュー d には満足度の得点スコアがあり、観測される得点を y_d とおく。 y_d については、 $Q + 1$ 点尺度の整数の点が振られているとする。最低点を 0 点とくと、 $y_d \in \{0, 1, \dots, Q\}$ となる。

3.2 Text Allocation Model: Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) は、Blei and Jordan (2003) によって提示されたモデルである。各文書に登場する語 w_i が以下のような V 次元のカテゴリカル分布に従っていると仮定するモデルである。

$$w_i \sim \text{Categorical}_V(\tilde{\phi}_i) \quad (1)$$

パラメータ $\tilde{\phi}_i$ については潜在的に K 種類存在すると仮定しており、この K 種類の潜在的な状態をトピックと呼ぶ。このとき、語 i が所属するトピックを示すパラメータを z_i とおく。 z_i は K 次元のベクトルであり第 i 語がトピック k に所属しているときは $z_{ik} = 1$ をとり、 $k' \neq k$ について $z_{ik'} = 0$ となる変数である。これを組み合わせてパラメータの要素を以下のように定義する。

$$\tilde{\phi}_{iv} = \prod_{k=1}^K \phi_{kv}^{z_{ik}} \quad (2)$$

また、 z_{ik} については、以下のようなカテゴリカル分布に従っていると仮定する。
ここで、 x_i は上で定義した文書の所属を示す変数である。

$$z_i \sim \text{Categorical}_K(\tilde{\theta}_i), \quad \tilde{\theta}_{ik} = \prod_{d=1}^D \theta_{dk}^{x_{id}} \quad (3)$$

加えて、パラメータ ϕ_k および θ_d については、カテゴリカル分布の共役事前分布であるディリクレ分布に従うと仮定し、 $\phi_k \sim \text{Dirichlet}(\beta)$, $\theta_d \sim \text{Dirichlet}(\alpha)$ とおく。

3.3 Binomial Factor Model (BFM)

LDA は自然言語の情報のみを分類対象としたモデルであるが⁸、Binomial Factor Model (BFM) は得点情報のみを分類対象にしたモデルである。まず、レビュー d に付されている得点について以下の分布を仮定する。

$$y_d \sim \text{Binomial}(Q, \tilde{\psi}_d) \quad (4)$$

ここで、 $\tilde{\psi}_d$ について、LDA と同様に K 個の潜在的な状態 (トピック) を仮定する。所属する状態を示す変数を c_d とおく。 c_d は K 次元のベクトルパラメータであり、第 k ビックに所属しているとき $c_{dk} = 1$ となり、 $k' \neq k$ について $c_{dk'} = 0$ である。

$$\tilde{\psi}_d = \prod_{k=1}^K \psi_k^{c_{dk}} \quad (5)$$

また、これも LDA と同様に、 c_d についても以下のような分布を仮定する。

$$c_d \sim \text{Categorical}_K(\theta_d) \quad (6)$$

パラメータ ψ_{kj} については、二項分布の共役事前分布であるベータ分布に従うと仮定する。また、パラメータ θ_d についてはカテゴリカル分布の共役事前分布であるディリクレ分布に従うと仮定する。それぞれ $\psi_{kj} \sim \text{Beta}(\gamma)$ 、 $\theta_d \sim \text{Dirichlet}(\alpha)$ とおく。完全条件付き事後分布は以下のようになる。

$$\pi(c, \Psi, \theta | y) \propto \pi(y | c, \Psi) \pi(c | \theta) \pi(\Psi) \pi(\theta) \quad (7)$$

ここで、LDA と同様に collapsed gibbs sampling によって c_d の簡易な事後分布を

得ることができる。

$$\pi(c|y) \propto \int \int \pi(c, \Psi, \theta|y) d\Psi d\theta \quad (8)$$

得点の期待値は $Q \times \psi_k$ となるため、この値が大きいほど高い得点を得られる傾向が高いことが示唆される。

3.4 提案モデル: Score-Text Allocation Model (STAM)

LDA と BFM を合成させて、共通のトピックを仮定するモデル STAM (Score-text Allocation Model) を定義する。二つの異なるデータを考慮するモデルとして、Mimno et al. (2009) においては Polylingual Topic Model (または Joint Topic Model) を提示しているが、本研究においてもレビューとスコアの2つのデータを考慮している。まずレビュー d のスコアについて、以下のように二項分布に従うとして定義する。このとき、パラメータは潜在的に存在するトピックに依存し、その所属を示すパラメータを c_{dk} とおく。

$$y_d \sim \text{Binomial}(Q, \tilde{\psi}_d), \quad \tilde{\psi}_d = \prod_{k=1}^K \psi_k^{c_{dk}} \quad (9)$$

また、レビュー観測される語 w_i についても LDA と同様の仮定をおく。

$$w_i \sim \text{Categorical}_V(\tilde{\phi}_i), \quad \tilde{\phi}_{iv} = \prod_{k=1}^K \phi_{kv}^{z_{ik}} \quad (10)$$

モデルには潜在的なパラメータとして z_i と c_d の2つを含むことになる。ここで、 z_i について、語 i が文書 d に所属していれば、 c_d と同じパラメータ θ_d を共有するという仮定をおく。

$$z_i \sim \text{Categorical}_K(\tilde{\theta}_i), \quad \tilde{\theta}_{ik} = \prod_{d=1}^D \theta_{dk}^{x_{id}} \quad (11)$$

$$c_d \sim \text{Categorical}_K(\theta_d) \quad (12)$$

パラメータの事前分布について、 ϕ_i および θ_d についてはディリクレ分布を仮定し、 ψ_k については、二項分布と共役なベータ分布に従うと仮定する。それぞれ $\phi_i \sim \text{Dirichlet}(\beta)$ 、 $\theta_d \sim \text{Dirichlet}(\alpha)$ 、 $\psi_k \sim \text{Beta}(\gamma)$ とおく。LDA と BFM を組み合わせることで STAM が得られている。パラメータ θ が共有され、ここから潜在的なトピックアロケーションパラメータ c および z が生成される。

4 データ

4.1 データの収集

本研究では、日本の価格比較サイトである Kakaku.com (<https://kakaku.com>) から収集したレビューデータを分析の対象としている。Kakaku.com は、Kakaku.com, Inc. によって運営されており、主に電化製品の実売価格比較及び製品レビューの閲覧が可能なプラットフォームである。このサイトは、主に日本市場で販売されている製品の情報を扱っており、レビューも日本語で書かれている。

研究の対象製品カテゴリーとしては、スマートフォンのメーカーの中で Apple Inc. によって製造・販売されている iPhone シリーズを選定した。分析期間は、日本における iPhone 3G の発売開始である2008年8月から2018年2月までの約9年半と設定している。この期間に発売された iPhone の各モデルに関するレビューを分析の対象としている。データの収集方法としては、Python を用いたウェブスクレイピングが行われた。収集されたデータには、レビューの投稿日時、製品名、レビュータイトル、レビュー本文、投稿者名、そして得点が含まれている。得点はレビューの総合得点を

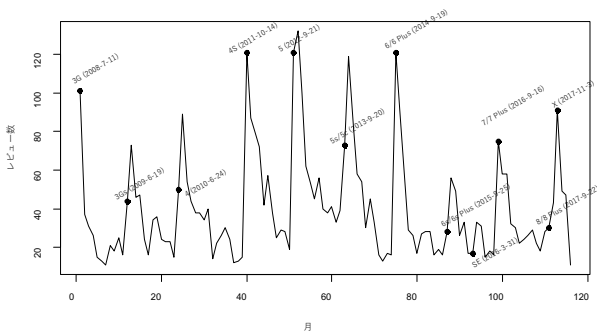


図1 レビュー数の推移

表す「満足度」を対象としており、1 から 5 点の整数値で評価されている。

分析対象となったレビューは、本文が存在し、評価得点に欠損がない4,684件である。レビュー件数を月別に集計し、それぞれの iPhone モデルの発売月と合わせて図示した結果(1)、新モデルの発売月にはレビューの件数が顕著に増加していることが確認された。また、分析期間全体を通じて、一定数のレビューが継続的に投稿されていることも明らかになった。

4.2 データの前処理

分析対象となったレビューの本文について、日本語で記述されているため、日本語用の形態素解析器を用いる。本研究では、Kudo et al. (2004) によって開発された汎用形態素解析エンジンである MeCab を用いて単語を分解している。また、単語を分解するための辞書として Sato et al. (2017) によって開発、管理されている新語辞書である ipadic-Neologd を利用している。これによって単語を分解し、名詞、動詞、形容詞を抽出する。結果として11,790個の語彙によって構成された354,760 個の語が分析対象となる。こうして得られた分析対象は、文書数 $D = 4,684$ 、各評価観点について $Q - 1 = 5$ 、語彙数 $V = 11,790$ 、コーパス長 $N = 354,760$ のデータセットである。評価データ y は $0, \dots, Q$ の値をとるベクトルであり、自然言語 w は $N \times V$ の $\{0, 1\}$ の2値を取る行列である。また、文書インデックス x は $N \times D$ の $\{0, 1\}$ の2値をとる行列となる。

5 分析結果

5.1 トピック数 K の決定

まずは適合度の比較のためにいくつかの指標を計算する。第1は、インサンブルの適合度を比較する指標として対数周辺尤度を計算する。本研究では、Newton and Raftery (1994) に従い、調和平均で周辺尤度を得る。この方法は Griffiths and Steyvers (2004) においても最適なトピック数を検討するためにも用いられている。

図2はトピック数とモデルの周辺尤度との関係である。周辺尤度は値が大きいほど適合がよいといえる指標である。ここから、 $K = 50$ 周辺までは急激に適合度が上昇し、 $K = 100$ のときにもっとも適合度が高くなり、それ以上になると緩やかに適合度が低

くなるという傾向が見て取れる。したがって、周辺尤度の観点からは $K = 100$ がもっとも適切なモデルであるといえる。したがって、以降では $K = 100$ のモデルをについて考察してく。

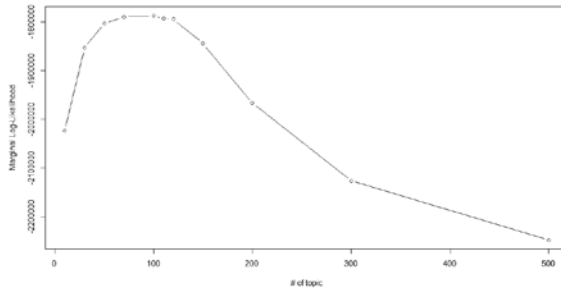


図2 トピック数 K の決定における Marginal Log-Likelihood

5.2 トピックの解釈

次に、モデルから得られたトピックの分析結果について考察する。LDA をはじめとした通常のトピックモデルにおいては、分析結果の解釈において、トピックごとに所属する確率の高い語を確認する必要がある。さらに、そのトピックが「良い」話題についてまとめられたものであるのか、「悪い」話題についてまとめられたものであるのかについても語から検討しなければならない。しかしながら、本研究の分析モデルにおいては、指針としてレビューのスコアがあり、トピックごとにレビューのスコアの期待値に影響を与えるパラメータ ψ_k が推定されている。この ψ_k を確認することで、当該トピックが良いレビューについてまとめられたものなのか悪いレビューについてまとめられたものなのか、数値で理解することができる。これによって、それぞれのトピックの所属確率の高い単語が良いレビューと関係しているのか悪いレビューと関係しているのかも考察することができる。

表1は全てのトピックの中から、 ψ_k の値が上位5位以内か、下位5位以内にあるトピックを挙げたものである。また、各トピックの中で出現確率を表す ϕ_{kv} が高い10個の単語を示している。 ψ_k は該当のトピックが現れるレビュー文は高い満足度を示すことを意味しているため、言葉と満足度の関係性を一目にわかることができる。例

えば、上位5位以内のトピックからみると、満足度の高いレビューには具体的な表現より、感情表現や抽象的な言葉が多く入っていることに対して、下位5位以内のトピックにはより具体的な不満の対象が出現されていることがわかる。

表1 ψ_k の上位5位及び下位5位以内のトピック

Rank	ψ_k	Topic #	Top 10 Words (ordered by ϕ_{kv})
1	0.998	68	感, 美しい, さ, 良さ, 所有, 操作, 抜群, 重量, コンパクト, 質感
2	0.998	57	問題, 特に, ない, 全区, バッテリー, 全然, 様子見, やっぱ, スリム, お世話
3	0.996	46	使い, 易い, 良い, 安さ, 辛い, 心地, 影響, 見易い, 未だ, html
4	0.996	99	とても, サクサク, やすい, バッテリー, 綺麗, 初めて, 勢い, 初心者, しっかり, スマート
5	0.996	32	満足, 大変, 概ね, 総じて, galaxynote, 不都合, おおむね, 懐かしい, 口コミ, 重宝
...			
96	0.194	1	悪い, 電池, 持ち, 最悪, もち, 全然, 消耗, マイナス, すぐ, むしろ
97	0.134	67	悪い, 追記, 良い, だいふ, そこそこ, 評判, 微妙, ふつう, 中古, 特別
98	0.080	34	必要, 接続, 登録, 移行, ダウンロード, 作業, 終了, 当方, 苦勞, インストール
99	0.031	11	対応, 選択, 意味, 数, hz, 恩恵, 的, 現時点, 帯, 倍
100	0.002	43	通話, 最初, 品質, 声, 相手, 初めて, 時々, はっきり, どう, 間

トピックの解釈を容易にするため、各トピックから ϕ_{kv} (トピックの単語生成確率) が高い単語に基づいて各トピックを定義する。表2はトピックの一部を示したものであり、Kakaku.comの評価項目に基づき、各トピックを定義した上で、関連のある上位概念で分類を行ったものの例である。

表2 トピックの出現単語による定義と分類

分類	関連項目	Topic #	Top 5 Words (ordered by ϕ_{kv})	ψ_k
基本性能	レスポンス	4	反応, 起動, 早い, タッチ, 遅い	0.838
	通話音質	43	通話, 最初, 品質, 声, 相手	0.002
	画面表示	71	ゲーム, 結構, 凄いい, 綺麗, でかい	0.953
	通話音質	92	電波, 場所, 圏外, 状況, 状態	0.711
	バッテリー	1	悪い, 電池, 持ち, 最悪, もち	0.194
	バッテリー	45	バッテリー, 早い, 充電, ケーブル, 常に	0.954
デザイン	デザイン	68	感, 美しい, さ, 良さ, 所有	0.998
	デザイン	100	色, ゴールド, 黒, ブラック, 白	0.984
	デザイン	18	ガラス, 傷, 背面, 質感, アルミ	0.894
	携帯性	21	薄い, 軽い, 速い, 長い, バッテリー	0.963
	携帯性	26	ポケット, 画面, 胸, 大きい, 丸み	0.991
	携帯性	70	操作, サイズ, 片手, 大きい, 手	0.818
機能	カメラ	52	カメラ, 写真, 撮影, 画質, 動画	0.830
	呼出音・音楽	62	音楽, 再生, 動画, 付属, イイ	0.880
	呼出音・音楽	69	スピーカー, 音質, クリア, 音量, 音	0.891

ここでは例として外観や性能に関するトピックを選別して、それらのパターンを確認するため図示する。図3と図4である。図3はトピックの肯定・否定(相対)的トピックの比率の変化を示し、図4は文章内各トピックの比率の変化を示している。

2008年から2018年にかけての iPhone に関するレビュートレンドを示す図3は、市場成長期と消費者期待の変遷を映し出している。2008年から2013年にかけてのポジティブなレビューの上昇は、iPhone が技術革新を遂げていた期間と重なる。この時期、Apple は画期的な新モデルを次々に発表し、それぞれが前モデルを大きく上回る性能を有していたため、消費者の期待は高まり、肯定的なレビューが増加した。Tellis and Johnson (2007) の研究によると、市場の初期成長段階では、消費者の楽観主義と製品への好意的な評価が伴うことが多い。

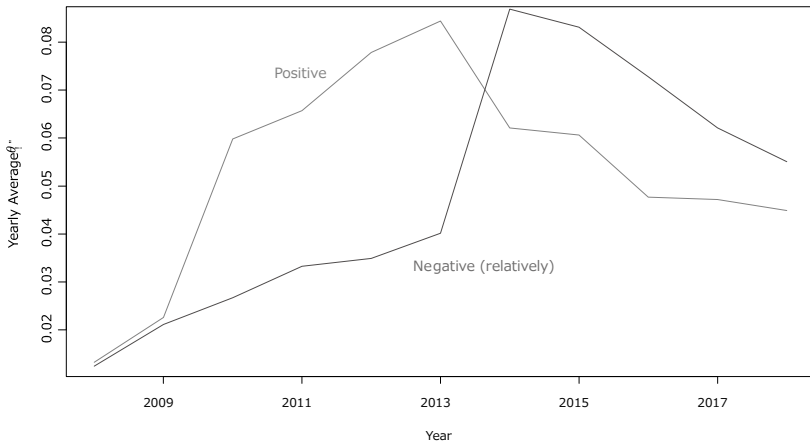


図3 トピックの出現パターン：肯定及び否定トピックの推移

しかし、2015年にポジティブなレビューがピークに達した後の下降は、市場の成熟と消費者期待の変化を示唆している。スマートフォン市場が飽和状態に近づき、技術進歩が段階的なものになるにつれて、消費者の期待もそれに合わせて変化した。Rogers のイノベーションの普及理論によると、製品が市場の飽和に達すると、新規採用者の増加率が鈍化し、製品への驚きの感覚が薄れるという。消費者は iPhone を先行モデルだけでなく、同等あるいはそれ以上の機能を低価格で提供する競合他社と比較するようになったのである。

2011年から2015年にかけてのネガティブなレビューの急増は、消費者期待に応えられなかった特定の製品リリースや、「アンテナゲート」や「バンドゲート」のような問題に関連している可能性がある。Clemons (2008) によれば、製品の失敗は消費者信頼に重大な影響を与え、否定的な認識を増大させる。2015年以降のネガティブなレビューの減少は、Apple が品質問題に対処し、消費者のフィードバックに応じて対応策を講じた結果、あるいは消費者が不満を他のフォーラムで表現するようになった結果かもしれない。また、スマートフォンの改善が段階的なものとなる成熟市場で、消費者がより現実的な期待を持つようになった反映である可能性もある。

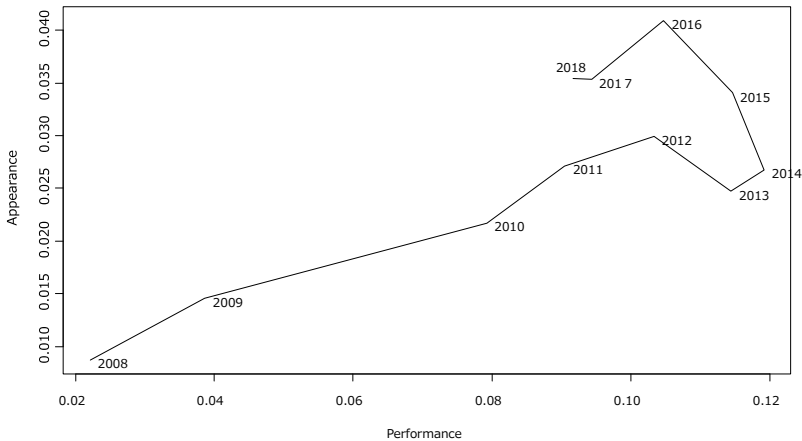


図4 トピックの出現パターン：デザイン及び性能トピックの推移

図4はiPhoneに対してレビューで言及されたトピックの割合の変化を示している。横軸が性能を、縦軸が外観およびデザインのトピックを表しており、時間経過に伴う変化を捉えている。2007年のiPhoneの登場以来、市場は急速な成長を遂げ、消費者の期待も変遷してきた。初期の段階では、革新的な機能とユーザーインターフェースが高く評価され、性能に関する言及が多かった。しかし、市場が成熟し競争が激化すると、消費者はより多様な機能とデザインに注目を移し始めた。図4によれば、時間が経つにつれて性能よりも外観やデザインのトピックの言及割合が増加している。これは、性能の進化が一定の水準に達し、消費者が製品選択の際に外観やデザインを重視するようになったことを示唆している可能性がある。また、市場の成熟とともに、

製品の差別化要因としてデザインの重要性が増していることが確認されている。

具体的には、性能の向上が消費者の基本的な要求を満たすレベルに達した後、デザインは消費者の情緒的、社会的ニーズに訴える要素としての役割を担うようになったと考えられる。これは、製品の形状、色、素材などの外観が、個人のアイデンティティやステータスの象徴として機能し始めたことを意味している。したがって、製品選択の決定要因として性能だけでなくデザインも同様に重要であることが、図4から読み取れる。

市場の成長期においては、消費者の期待は常に変化している。iPhone に関するこの図からは、性能と外観のトピックの重要性が時とともにどのように変わっていったかが見て取れる。消費者が製品に求める価値がどのように進化しているかを理解することは、市場戦略を立てる上で非常に重要である。

6 結論と今後の課題

本研究では、広く用いられるトピックモデル LDA を発展させた新しいモデルを開発し、自然言語で記述されたテキスト情報と数値の評価情報を併せ持つオンラインレビューに適用した。これにより、テキストとスコア情報を同時に分類することが可能となった。実証分析においては、日本市場での iPhone3G から iPhoneX に至る10年間のレビューデータに対して提案モデルを適用し、以下の結果を得た。LDA と比較して、提案モデルはマーケティングの意思決定においてより有用な情報を提供する。LDA では文書の感情やムードを直接解釈することは困難であるが、本モデルでは各トピックの平均評価スコアを算出するため、トピックが良い評価に関連しているのか、悪い評価に関連しているのかを直接判定できる。また、モデルの出力に基づき、トピックと製品の側面との関連性を視覚的に解釈することが可能であり、評価スコアだけでは見出せないスコアに影響を与える要因を発見する手がかりを提供する。

今後の研究の方向性としては、トピック数の問題点が挙げられる。分析者が解釈可能な範囲内でトピック数を設定することがマーケティングの意思決定において重要であり、 $K = 100$ のモデルでは全トピックを解釈することに困難が伴う。将来的には、パチンコ割り当てモデルのような階層モデルを用いることで、この問題に対処することが考えられる。また、実際の適用範囲として、本研究で扱ったスマートフォンのオ

ンラインレビューに限定せず、他のスマートフォンについても分析を行い、提案モデルの特性をさらに検証する必要がある。

参考文献

- Anderson, E. T. and D. I. Simester (2014) “Reviews without a purchase: Low ratings, loyal customers, and deception,” *Journal of Marketing Research*, **51**(3), pp. 249-269.
- Balducci, B. and D. Marinova (2018) “Unstructured data in marketing,” *Journal of the Academy of Marketing Science*, **46**(4), pp. 557-590.
- Berger, J., A. Humphreys, S. Ludwig, W. W. Moe, O. Netzer, and D. A. Schweidel (2019) “Uniting the tribes: Using text for marketing insight,” *Journal of Marketing*, **84**(1), pp. 1-25.
- Berger, J. and K. L. Milkman (2012) “What makes online content viral?” *Journal of Marketing Research*, **49**(2), pp. 192-205.
- Blei, D. M. and M. I. Jordan (2003) “Modeling annotated data,” in *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*, pp. 127-134: Association for Computing Machinery.
- Blei, D. M. and J. D. Lafferty (2006) “Dynamic topic models,” in *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pp. 113-120, New York, NY, USA: Association for Computing Machinery.
- Büschken, J. and G. M. Allenby (2016) “Sentence-based text analysis for customer reviews,” *Marketing Science*, **35**(6), pp. 953-975.
- Chen, Z. and N. H. Lurie (2013) “Temporal contiguity and negativity bias in the impact of online word of mouth,” *Journal of Marketing Research*, **50**(4), pp. 463-476.
- Chevalier, J. A. and D. Mayzlin (2006) “The effect of word of mouth on sales: Online book reviews,” *Journal of Marketing Research*, **43**(3), pp. 345-354.
- Clemons, E. K. (2008) “How information changes consumer behavior and how consumer behavior determines corporate strategy,” *Journal of Management Information Systems*, **25**(2), pp. 13-40.
- Griffiths, T. L. and M. Steyvers (2004) “Finding scientific topics,” *Proceedings of the National Academy of Sciences*, **101**(suppl 1), pp. 5228-5235.
- Hewett, K., W. Rand, R. T. Rust, and H. J. Van Heerde (2016) “Brand buzz in the echoverse,” *Journal of Marketing*, **80**(3), pp. 1-24.
- Humphreys, A. and R. J. Wang (2018) “Automated text analysis for consumer research,” *Journal of Consumer Research*, **44**(6), pp. 1274-1306.
- Kannan, P. K. and H. A. Li (2017) “Digital marketing: A framework, review and research agenda,” *International Journal of Research in Marketing*, **34**(1), pp. 22-45.
- Kudo, T., K. Yamamoto, and Y. Matsumoto (2004) “Applying conditional random fields to japanese morphological analysis,” in *Proceedings of the 2004 conference on empirical*

- methods in natural language processing*, pp. 230–237: Association for Computational Linguistics.
- Lamberton, C. and A. T. Stephen (2016) “A thematic exploration of digital, social media, and mobile marketing: research evolution from 2000 to 2015 and an agenda for future inquiry,” *Journal of Marketing*, **80**(6), pp. 146–172.
- Lee, L. and E. T. Bradlow (2011) “Modeling the internet and the web: Probabilistic methods and algorithms,” *Journal of Marketing Research*, **48**(4), pp. 733–748.
- Liu, X., A. C. Burns, and Y. Hou (2017) “An investigation of brand-related user-generated content on twitter,” *Journal of Advertising*, **46**(2), pp. 236–247.
- Liu, Y. (2006) “Word of mouth for movies: Its dynamics and impact on box office revenue,” *Journal of Marketing*, **70**(3), pp. 74–89.
- Luo, X. (2009) “Quantifying the long-term impact of negative word of mouth on cash flows and stock prices,” *Marketing Science*, **28**(1), pp. 148–165, Accessed 15 Feb. 2024.
- Mimno, D., H. M. Wallach, J. Naradowsky, D. A. Smith, and A. McCallum (2009) “Polylingual topic models,” in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pp. 880–889: Association for Computational Linguistics.
- Nam, H., Y. V. Joshi, and P. K. Kannan (2017) “Harvesting brand information from social tags,” *Journal of Marketing*, **81**(4), pp. 88–108.
- Newton, M. A. and A. E. Raftery (1994) “Approximate bayesian inference with the weighted likelihood bootstrap,” *Journal of the Royal Statistical Society. Series B (Methodological)*, **56**(1), pp. 3–48.
- Onishi, H. and P. Manchanda (2012) “Marketing activity, blogging and sales,” *International Journal of Research in Marketing*, **29**(3), pp. 221–234.
- Pang, B. and L. Lee (2008) *Opinion Mining and Sentiment Analysis*, **2**, pp.1–135.
- Rosen-Zvi, M., T. Griffiths, M. Steyvers, and P. Smyth (2004) “The author-topic model for authors and documents,” in *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, UAI ’04, pp. 487–494, Arlington, Virginia, USA: AUAI Press.
- Sato, T., T. Hashimoto, and M. Okumura (2017) “Implementation of a word segmentation dictionary called mecab-ipadic-neologd and study on how to use it effectively for information retrieval,” in *Proceedings of the Twenty-Three Annual Meeting of the Association for Natural Language Processing*, The Association for Natural Language Processing.
- Stephen, A. T. and J. Galak (2012) “The effects of traditional and social earned media on sales: A study of a microlending marketplace,” *Journal of Marketing Research*, **49**(5), pp. 624–639.
- Tellis, G. J. and J. Johnson (2007) “The value of quality,” *Marketing Science*, **26**(6), pp. 758–773, Accessed 15 Feb. 2024.
- Tirunillai, S. and G. J. Tellis (2014) “Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation,” *Journal of*

Marketing Research, **51**(4), pp. 463–479.

- (2012) “Does chatter really matter? dynamics of user-generated content and stock performance,” *Marketing Science*, **31**(2), pp. 198–215, Accessed 15 Feb. 2024.
- Toubia, O., G. Iyengar, R. Bunnell, and A. Lemaire (2019) “Extracting features of entertainment products: A guided latent dirichlet allocation approach informed by the psychology of media consumption,” *Journal of Marketing Research*, **56**(1), pp. 18–36.
- Trusov, M., L. Ma, and Z. Jamal (2016) “Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting,” *Marketing Science*, **35**(3), pp. 405–426.

