

Title	機械学習とマーケティング
Sub Title	Machine learning and marketing
Author	里村, 卓也(Satomura, Takuya)
Publisher	慶應義塾大学出版会
Publication year	2020
Jtitle	三田商学研究 (Mita business review). Vol.63, No.4 (2020. 10) ,p.111- 121
JaLC DOI	
Abstract	近年, マーケティング分野でのビッグデータの活用が活発になってきている。これに対応して機械学習のアルゴリズムを利用した計量的マーケティングの研究も進みつつある。本稿では, マーケティング研究と機械学習の関係について概観し, また機械学習のアルゴリズムを利用したマーケティング分析に関する研究を紹介する。さらに機械学習とマーケティング研究の今後について展望する。
Notes	堀越比呂志教授退任記念号 論文
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234698-20201000-0111

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

機械学習とマーケティング

里村卓也

<要約>

近年、マーケティング分野でのビッグデータの活用が活発になってきている。これに対応して機械学習のアルゴリズムを利用した計量的マーケティングの研究も進みつつある。本稿では、マーケティング研究と機械学習の関係について概観し、また機械学習のアルゴリズムを利用したマーケティング分析に関する研究を紹介する。さらに機械学習とマーケティング研究の今後について展望する。

<キーワード>

アルゴリズム, ガウス過程, 確率的最適化, ノンパラメトリックベイズ, 変分ベイズ法

1. はじめに

マーケティングにおいてビッグデータや機械学習の活用が注目されている。マーケティング実務でのビッグデータや機械学習の活用については Amazon や Netflix の例が有名であるが (Domingos 2015), マーケティング研究でもビッグデータや機械学習への関心はここ数年の間に高まってきている。Marketing Science 誌では2016年35巻3号において Marketing Science and Big Data (Chintagunta, Hanssens and Hauser 2016) として特集号が組まれた。また Wedel and Kannan (2016) ではビッグデータとマーケティング分析に関する議論が, Bradlow, Gangwar, Kopalle and Voleti (2017) ではビッグデータと小売業に関する議論がなされている。さらにマーケティングだけでなく、マーケティング分析と関連の深い経済学においても機械学習に対する関心が高まっている (Athey 2019)。

本論では、特に機械学習のアルゴリズムの発展がマーケティング分析の研究に及ぼす影響について考える。なお、ビッグデータや機械学習がマーケティング活動や消費者行動に及ぼす影響についての研究 (例えば Davenport, Guha, Grewal and Bressgott 2020や Shiller 2020) は、もちろん大き

な関心を集めている分野ではあるが、本稿では機械学習のアルゴリズムを中心にマーケティング研究との関連を見ていく。そのためにマーケティング分析における最近の機械学習のアルゴリズム利用についていくつかを取り上げ、さらに今後のマーケティング分析の研究の方向性を探る。

本稿の構成は次のとおりである。2. では機械学習とマーケティングの関係について述べ、3. 以降ではマーケティング分析における機械学習の適用について見ていく。3. ではマーケティング分析で利用されるベイズ推定と機械学習のアルゴリズムについて述べる。特に、計算コストを軽減するために有用な変分ベイズ法についてマーケティング研究での適用について概観する。4. では、分布を仮定しないノンパラメトリックベイズと、その一つであるガウス過程のマーケティング研究での適用について概観する。5. では上記以外の機械学習について、非構造化データ、スパース性への対応、その他の機械学習アルゴリズムのマーケティング研究での適用について述べる。6. では機械学習とマーケティング研究の今後について展望する。

2. 機械学習とマーケティング分析

2.1. 機械学習とは

機械学習は広義にも狭義にも定義できるが、本稿では狭義に考え、機械学習とは「データに適用されるようにデザインされたアルゴリズムであり、主な領域として予測（回帰）、判別、クラスタリングに焦点をあてたもの（Athey 2019）」としておく。機械学習には、大きく分けると「教師あり学習」と、「教師なし学習」がある。教師あり学習は、正解である出力データを与え、入力データからの出力が、正解である出力データに適合するように識別規則を得るものである。一方、教師なし学習は、正解となる出力データはなく、与えられた入力データのみから出力を決定する方法である。機械学習の適用領域である予測（回帰）と判別では教師あり学習が行われ、クラスタリングでは教師なし学習が行われる。

機械学習では、データにアルゴリズムを適用する際には、正解である出力データと学習による予測との差である訓練誤差を小さくするようにモデルの学習を行う。ただし、訓練誤差を小さくするだけでは、学習データにのみ適合することによる過学習が起こってしまう可能性があるため、未知のサンプルに対する汎化能力である汎化誤差についても考慮する必要がある。

ところで人間の活動である認識をアルゴリズムにより実現する工学的手法に、パターン認識がある。パターン認識とは「物事の類型を知るはたらき、およびその内容（平井 2012）」のことである。マーケティング分野での機械学習の利用目的は、データからマーケティング活動に有用なパターンを見つけ出そうというものである。機械学習ではパターン認識を行っているとして理解してよからう。

2.2. 統計モデルにとっての機械学習

機械学習の登場前から、データの分析のためには統計モデルが利用されてきた。マーケティング分析においても、回帰分析、分散分析、多項ロジットモデル、潜在クラスモデル、階層ベイズ

モデル、構造モデリング等、多くの統計モデルが利用されてきた (Wedel and Kannan 2016)。従来から利用されてきた統計モデルと機械学習については、両者とも主な用途は予測、判別、分類であり、両者の間に違いはない。実際のところ、機械学習として利用されているアルゴリズムの中には、従来から利用されてきた統計モデルも多い。ならば、従来から利用されてきた統計モデルと機械学習の違いは何であろうか。

Breiman (2001) によると統計モデルを利用したデータ分析には二つの文化がある。一つは伝統的なデータ・モデリング文化であり、データは特定の確率モデルから生成されたものであると考える。もう一つは新しいアルゴリズム・モデリング文化であり、データ生成のメカニズムは未知としてアルゴリズム・モデルを用いる。機械学習では主にアルゴリズム・モデリングが利用され、分析対象のメカニズムを解明することよりも、入力と出力の関係をメカニズムなしに直接関係づけるアルゴリズムに重点を置いている。

このように確率モデルでは、データの生成メカニズムに興味があるため生成モデルを出発点とし、対象のメカニズム (Why) に興味がある場合に利用される。これに対し、機械学習ではデータから直接出力を得るために、計算能力や処理速度に重点が置かれ、活用 (What, How) に興味がある場合に利用される。

なお、機械学習ではアルゴリズムによる学習結果を解釈することが一般的には難しく、この解釈可能性が利用のうえでの問題とされてきた。しかし、近年では機械学習の学習結果の解釈可能性に関する研究も進められており、解釈可能な機械学習 (Interpretable Machine Learning) として (Molnar 2019) 注目されている。

2.3. データとモデルの関係

確率モデルであれ機械学習であれ、データを目的に沿った形で利用するためには統計モデルが必要である。マーケティング分析においては、モデルの複雑さとデータサイズの間を考慮する必要がある。これに関して Wedel and Kannan (2016) では big stats on small data と small stats on big data という二つの立場があることを述べている。彼らによると、big stats on small data の立場では行動メカニズムと最適意思決定にもとづく厳密で完全な統計モデルを構築し、構造化された比較的小さいサイズのデータを用いて統計モデルの推定を行う。一方、small stats on big data の立場では探索あるいは記述目的で非構造化データからリアルタイムで情報を抽出し、計算能力やアルゴリズムに注力する。また、彼らによると、実務的研究では small stats on big data に、学術的研究では big stats on small data に焦点をあてている。

データの質に関する議論も重要である。Bradlow et al. (2017) では流通業におけるビッグデータ利用の可能性について論じているが、その中でデータの質の問題を挙げている。Bradlow et al. (2017) によると顧客個人レベルでのデータは good data (recent data) と bad data (old data) があり、小売業者が収集しているデータはビッグデータとなっても、過去のマーケティング施策と顧客行動の結果である old data であることが多い。加えて、彼らは、今後どうすべきかという意思決定には old data は利用価値が低いことを指摘している。

マーケティング研究においては、予測精度を高めるだけでなく、事象についての因果関係やメカニズムの解明も重要な目的となってくる。そのため、モデリングにおいても理論をもとにした厳密な議論が可能となるように、問題に応じてモデル構造を考える必要がある。このようにマーケティング分析においてデータやアルゴリズムを考える場合には、データのサイズ・質についての議論に加えて理論とモデルの関係についても検討しておく必要がある。

2.4. マーケティング分析における機械学習の利用

計量マーケティングのトップジャーナルである *Journal Marketing Research* 誌や *Marketing Science* 誌などにもここ数年、機械学習についての研究が多く発表されるようになってきた。これらの研究には、既存の機械学習のアルゴリズムを適用したものもあるが、機械学習のアルゴリズムを発展させ新しいマーケティング分析の方法を開発しようというものも多く見られる。

ビッグデータの利用はマーケティング分析において重要なテーマとなっているが、その際には、計算量の増大が大きな問題となる。機械学習分野で開発されたアルゴリズムは、この問題を解決してくれることが期待されており、新しい研究領域として活発に取り組みされている。

3. ベイズ推定と機械学習のアルゴリズム

3.1. ベイズ推定における問題点と対応

ベイズ推定では階層モデルを用いて、個人間の異質性を取り込むことが可能である。個人別のマーケティングが可能となった現代では、マーケティング課題に対して階層ベイズモデルの有用性が着目されてきた。一方、個人レベルでの異質性を考慮した階層ベイズモデルをビッグデータに適用する場合、特に共役でない事前分布を用いたモデルについて MCMC (Markov Chain Monte Carlo) 法でパラメータの事後分布を求める場合には計算コストが大きくなってしまいう問題がある。(Allenby, Bradlow, George, Liechty and McCulloch 2014)。

このような計算コストを軽減するための一つの解決方法は近似的手法を利用することである。代表的な近似の方法としては、尤度関数を近似することで閉形式を利用しやすくする方法がある。Bradlow, Hardie and Fader (2002) では負の二項分布を級数展開で近似することで事後分布を閉形式で表現できることを示している。また Everson and Bradlow (2002) ではベータ二項分布を、Miller, Bradlow and Dayaratna (2006) ではロジットモデルを、それぞれ級数展開で近似することで、事後分布を閉形式で表現し、計算コストを削減している。また尤度関数を近似するのではなく、事後分布を近似する方法がある。例えばロジスティック回帰モデルではラプラス近似を用いて事後分布を正規分布で表現する方法がある (Bishop 2006)。

3.2. 変分ベイズ法

一方、機械学習で用いられる変分ベイズ法は事後分布を因子分解可能な分布で近似し、変分下限を最大にすることで近似事後分布を求めるものである。変分法とは関数の関数(汎関数)を最

小化する関数を求める方法である。変分ベイズ法では、求める事後分布を因子分解可能な分布で近似する。そして関数をこの近似事後分布、汎関数を KL (Kullback-Leibler) 情報量とすると、KL 情報量を最小にする近似事後分布が求める分布となる。ここでサンプルからの対数周辺尤度についてイエンセンの不等式から求めた下限を変分下限と呼ぶことにする。すると対数周辺尤度は KL 情報量と変分下限の和になり、また対数周辺尤度は事後分布に関して定数であるので、KL 情報量を最小にする代わりに変分下限を最大にすることで、求める近似事後分布を得ることができる。これが変分ベイズ法である。変分ベイズ法は MCMC 法と異なり決定論的な手法である。Blei, Kucukelbir and McAuliffe (2017) によるとは、変分ベイズ法は近似的手法であるが大規模データセットおよび計算速度が必要な場合に向いている手法である。

変分ベイズ法では、変分下限の最大化のために、変分事後分布を微分して勾配法により変分事後分布のパラメータを求める。ただし、すべてのデータについて微分を行うことは計算負荷が高い。そこで、確率的最適化を適用し、データの一部をサンプリングして勾配法により最適化し、これを繰り返すことで最適値を得るのが確率変分ベイズ法 (Hoffman, Blei, Wang and Paisley 2013) である。確率的最適化とはデータを適切に分割して、ランダムに小さな部分問題を解いてゆき、結果的にデータ全体を用いた最適化を行う方法である (鈴木 2015)。また、変分下限を解析的に求めることが難しい場合に自動的に変分族を決定する自動微分変分法 (Kucukelbir, Ranganath, Gelman and Blei 2015) が提案されている。ただし自動微分変分法は勾配法による最適化を行うため離散確率変数を扱う確率モデルには利用できないという点に限界もある。

3.3. 変分ベイズ法のマーケティング分野での適用

Braun and McAuliffe (2010) ではビッグデータにおける離散選択モデルの推定方法として変分ベイズ法の利用を提案し、Ishigaki, Terui, Sato and Allenby (2018) もアイテム数が多いために個人レベルでは非常に疎なデータにおける選択モデルについて変分ベイズ法による推定を行っている。Braun and McAuliffe (2010) では MCMC 法で推定に 2 日間かかる問題が変分ベイズ法では 6 時間で推定できたと報告している。Ishigaki et al. (2018) では変分ベイズ法は MCMC 法より 20 から 50 倍効率的であったと報告している。Ansari, Li and Zhang (2018) では共変量付教師ありトピックモデルを提案し、12 万人による 900 万本の映画のランキングについて確率変分ベイズ法を適用している。Dzyabura and Hauser (2011) では変分ベイズ確率伝播アルゴリズムを利用してウェブ調査で適応的な質問をして消費者のヒューリスティックな意思決定ルールの推定を行う方法を提案している。

近年は Stan 言語 (Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, Guo, Li and Riddell 2017) のように MCMC 法や自動変分ベイズ法を利用できる確率的プログラミング言語も開発されており、分析者は確率モデルを指定するだけで MCMC 法や変分ベイズ法を利用することができる。このような確率的プログラミング言語を利用することで、マーケティング分野でビッグデータを扱う場合にもモデリング中心の研究が拡大することが期待される。

4. ノンパラメトリックベイズとガウス過程

4.1. ノンパラメトリックベイズ

マーケティングにおいてセグメンテーションを行う場合には、モデルベースの分析として潜在クラスモデルが利用される。潜在クラスモデルではセグメント数は事前に設定しておく必要があり、モデルの推定後に情報量基準や周辺尤度を用いてセグメント数の選択が行われる。これに対して、無限次元でのモデリングであるノンパラメトリックベイズを用いると事前にセグメント数を設定することなく、分析を行うことができる。ノンパラメトリックベイズモデルは有限混合モデルを無限次元化したものとみなすことができるため、無限混合モデルとも呼ばれている（佐藤 2016）。また、無限混合モデルを利用することで、クラスタリングだけでなく、時系列での構造変化や因子分析、スパースモデリングなどへ応用できる（佐藤 2016）。

4.2. ガウス過程

ノンパラメトリックベイズモデルの一つにガウス過程がある。ガウス過程は入力に対する出力関数が確率分布となる無限次元の正規分布であり、任意個数 N の入力点に対する出力が N 次元の多変量正規分布となる（持橋・大羽 2019）。このようにガウス過程は正規分布の一般化であると考えられる（Rasmussen and Williams 2006）。ガウス過程を利用することで、事前に次元数を決定することなく、データポイントの数に応じて正規分布を利用することができる。

ガウス過程の計算においても、データ数の増加に伴い計算量が増加するため、計算のためには工夫が必要である。ガウス過程の計算では補助変数法、変分ベイズ法、確率的勾配降下法を用いることでデータをミニバッチと呼ばれる小口サイズに分けて逐次的に学習する計算方法が使えるようになり、メモリー消費量と演算量を大幅に減らすことが可能である（持橋・大羽 2019）。

4.3. マーケティング分析におけるノンパラメトリックベイズの利用

マーケティング分析におけるノンパラメトリックベイズの利用としては、顧客間のインタラクションデータの潜在空間表現（Braun and Bonfrer 2011）、非線形ダイナミックモデルによる広告分析（Bruce 2019）、ランダムパーティションモデルを利用した製品のグループ化（Smith, Rossi and Allenby 2019）、ノンパラメトリックベイズ因子分析モデルによる動的広告反応の分析（Bruce, Becker and Reinartz 2020）などがある。

ガウス過程もマーケティング分野においても利用されている。Dew and Ansari (2018) ではロジットモデルの線形予測子において時間によって変化する効果を推定するために、これらの項の事前分布にガウス過程を利用している。また Dew, Ansari and Li (2020) では、個人間異質性と個人内異質性を表現するためガウス過程動的個人モデルを提案している。このモデルでは効用の個人間異質性だけでなく、時間による個人内異質性についてガウス過程を用いて表現している。

5. その他の機械学習のアルゴリズム

5.1. 非構造化マーケティングデータの分析

テキストや画像のデータのようにデータ構造として変数構造が決まっていないデータは非構造化データと呼ばれる。トピックモデル (Blei, Ng and Jordan 2003) はテキストデータの分析において利用されることが多いモデルである。トピックモデルでは文章の背後には潜在的なトピックがあり、その潜在的トピックをもとに単語が発生していると考えられる。

マーケティング分析でトピックモデルを利用した研究として、オンライン会話の分析 (Tirunillai and Tellis 2014) や規制の前後でのトピック比率の変化 (Puranam, Narayan and Kadiyali 2017) があるが、これらは Blei et al. (2003) の基本的なトピックモデルを利用している。一方、対象とする課題に応じてトピックモデルの拡張の研究も進んでいる。トピックモデルの拡張に関する研究にはセンテンス制約 LDA モデル (Büschken and Allenby 2016)、サーチ語とサーチ結果の 2 重 LDA モデル (Liu and Toubia 2018)、トピック比率が時間と共に変化する多重潜在変化点トピックモデル (Zhong and Schweidel 2020)、自己相関トピックモデル (Büschken and Allenby 2020) などがある。また、トピックモデルは生成モデルでもあるので、マーケティングでの応用に向いており、テキストデータの分析だけでなく、購買アイテムの分析 (Jacobs, Donkers and Fok 2016; Ishigaki et al. 2018) などにも利用されている。

5.2. スパース性にもとづくマーケティングデータの分析

多くの変数のうち、ほとんどはゼロで、ごく一部だけが非ゼロの値をとることを、スパース性と呼ぶ (富岡 2015)。変数が多くなるほど得られるデータがゼロの値をとることが多くなり、あるいは目的変数に影響しない説明変数が多くなる。このような問題に対して、従来は SSVS (Stochastic Search Variable Selection) を利用することで (例えば Gilbride, Allenby and Brazell 2006)、スパース性への対応を行ってきた。

それに対し、機械学習のアルゴリズムでは正則化法により変数選択を行いスパース性への対応を行う。正則化法では、罰則項を用いることで変数選択と回帰係数の推定を同時に実行する。この時、罰則項として L1 ノルムを用いる L1 正則化が LASSO (Least Absolute Shrinkage and Selection Operator) (Tibshirani 1996) である。ただし、LASSO では変数間に多重共線性があった場合には、それらの変数をまとめて選択するのではなく、ランダムに一つの変数を選択してしまう。このために効果があるのに変数が選択されないという結果を得る可能性がある。これではマーケティングの意思決定や予算配分で問題を引きおこす。そこで罰則項に L1 だけでなく L2 ノルムの罰則項を加えて L1 正則化の条件を緩めたものが Elastic Net (Zou and Hastie 2005) である。Elastic Net では、変数選択と同時に LASSO では不可能であった多重共線性の問題を解決することが可能となる。

Rutz, Trusov and Bucklin (2011) では Elastic Net を用いて、ペイドサーチ広告で使用された

キーワードのオンライン店舗訪問への間接効果を分析した。Xia, Chatterjee and May (2019) では複数カテゴリーでの購買を表現するために、機械学習のアルゴリズムである RBM (Restricted Boltzmann Machine) を利用し、RBM の学習に Elastic Net を利用した。これらの研究はビッグデータであっても消費者レベルでのキーワードや週別カテゴリー別のデータは多くがゼロであることに対応するものである。

正則化の方法は、このようなビッグデータへの対応以外にも、消費者間の異質性を考慮したモデルでも利用可能である。Chen, Iyengar and Iyengar (2017) ではコンジョイント分析において、セグメント別の効用が混合した多峰型の効用関数を利用するために、従来から利用されてきた有限混合分布モデルではなく正則化を用いた手法を提案している。

5.3. その他の機械学習アルゴリズムの利用

その他の機械学習をマーケティングに適用した研究では、機械学習のアルゴリズムを計算に利用するのではなく、アルゴリズムのアイデアを用いて消費者の意思決定を表現する研究もある。Lin, Zhang and Hauser (2015) では、消費者のヒューリスティクスに着目することで、この最適化問題を機械学習分野で用いられる多腕バンディット問題に類似した問題として帰着させ消費者の先見的行動を表現している。このように消費者の意思決定問題を機械学習のアルゴリズムに帰着させる視点は非常に興味深く、今後も発展が期待される。

また、機械学習ではエージェントが環境からの報酬を得ることで最適な行動を選択する強化学習と呼ばれる手法がある。これとは逆に、エージェントの行動は最適だと仮定して、これから報酬を学習するのが逆強化学習 (Ng and Russell 2000) である。ここで逆強化学習における報酬を効用と考え、逆強化学習は経済学やマーケティングで利用される構造推定と類似している (Athey 2019)。Halperin (2017) では forward-looking な消費者の選択モデルの構造推定で逆強化学習を利用することで、MCMC 法を利用せずにパラメータの推定を行っている。行動から報酬 (効用) を推定することは、マーケティング分析での主な関心であるため、これまでのマーケティング研究と機械学習をつなぐ研究であるといえよう。

3.2. で取り上げた確率的最適化もビッグデータに対応して機械学習で利用されるアルゴリズムである。確率的変分ベイズ法以外に、マーケティング分析では Gabel, Guhl and Klapper (2019) や Ruiz, Athey and Blei (2020) がショッピングバスケットデータの分析で確率的勾配降下法を利用してモデル学習を行い、商品間の関連を表現するマップを作成している。

これら以外にも、サポートベクターマシン (Cui and Curry 2005; Huang and Luo 2016)、ランダムフォレスト (Schwartz, Bradlow and Fader 2014) などの機械学習のアルゴリズムがマーケティング分析の研究で利用されている。

6. おわりに

伝統的にマーケティング研究では、マーケティング理論をもとにした分析モデルを利用して、

新しい知識の獲得を行ってきた。例えば、需要モデルにおいては、消費者の効用最大化に加えて、学習や参照価格、考慮集合、選択ヒューリスティクスなどの理論を数理的に表現することで、消費者の行動を解明してきた。このようにしてマーケティングにとっての bespoke model を構築することで、消費者や市場のメカニズムを理解する取り組みがなされてきた。

一方、マーケティングにおいてもビッグデータを分析することで得られる知見への関心の高まりと共に、テキストデータやイメージデータなどの非構造化データのように従来の方法では利用されなかったデータの活用にも関心が高まっている。

そこで、このようなビッグデータを活用するために、マーケティング研究においても、機械学習の利用が広まってきた。ここ数年間の機械学習を利用した研究数の増大は顕著である。1990年代中盤以降はマーケティング研究においてベイズ統計学の利用が広範囲に普及したが、2010年代中盤以降は特に計量的マーケティング研究において、機械学習の利用は重要なテーマの一つとなっている。ただし、本稿で見てきたように、機械学習のアルゴリズムをそのままデータに適用するのではなく、これまでのマーケティングや消費者行動の研究の中で蓄積されてきた理論や方法論も踏まえながら、両者を融合する形で発展が続いている。

機械学習に関する研究は、多くの関連分野を巻き込みながら、非常に速いスピードで進められている。マーケティング研究としては、このような機械学習のアルゴリズムの中から、マーケティング分野に固有の問題に利用可能なアルゴリズムを取り込んで問題を解決していくことはもちろんのこと、マーケティング分野から機械学習のアルゴリズムの開発を促すような新しい研究のテーマを提供していくこともできれば、さらに両者の発展につながると考える。

参 考 文 献

- Allenby, G.M., E.T. Bradlow, E.I. George, J. Liechty and R.E. McCulloch (2014) "Perspectives on Bayesian Methods and Big Data," *Customer Needs and Solutions*, 1 (3), pp.169-175.
- Ansari, A., Y. Li and J.Z. Zhang (2018) "Probabilistic Topic Model for Hybrid Recommender Systems: A Stochastic Variational Bayesian Approach," *Marketing Science*, 37 (6), pp.987-1008.
- Athey, S. (2019) "The Impact of Machine Learning on Economics," in *The Economics of Artificial Intelligence: An Agenda*, A. Agrawal, J. Gans and A. Goldfarb eds., University of Chicago Press, pp. 507-552.
- Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*, Springer.
- Blei, D.M., A. Kucukelbir and J. McAuliffe (2017) "Variational Inference: A Review for Statisticians," *Journal of the American Statistical Association*, 112, pp.859-877.
- Blei, D.M., A.Y. Ng and M.I. Jordan (2003) "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, pp.993-1022.
- Bradlow, E.T., M. Gangwar, P. Kopalle and S. Voleti (2017) "The Role of Big Data and Predictive Analytics in Retailing," *Journal of Retailing*, 93 (1), pp.79-95.
- Bradlow, E.T., B.G.S. Hardie and P.S. Fader (2002) "Bayesian Inference for the Negative Binomial Distribution via Polynomial Expansions," *Journal of Computational and Graphical Statistics*, 11 (1), pp.189-201.
- Braun, M. and A. Bonfrer (2011) "Scalable Inference of Customer Similarities from Interactions Data Using Dirichlet Processes," *Marketing Science*, 30 (3), pp.513-531.
- Braun, M. and J. McAuliffe (2010) "Variational Inference for Large-Scale Models of Discrete Choice," *Journal of the American Statistical Association*, 105 (489), pp.324-335.

- Breiman, L. (2001) "Statistical Modeling: The Two Cultures," *Statistical Science*, 16(3), pp.199–215.
- Bruce, N.I. (2019) "Bayesian Nonparametric Dynamic Methods: Applications to Linear and Nonlinear Advertising Models," *Journal of Marketing Research*, 56(2), pp.211–229.
- Bruce, N.I., M. Becker and W. Reinartz (2020) "Communicating Brands in Television Advertising," *Journal of Marketing Research*, 57(2), pp.236–256.
- Büschken, J. and G.M. Allenby (2016) "Sentence-Based Text Analysis for Customer Reviews," *Marketing Science*, 35(6), pp.831–998.
- Büschken, J. and G.M. Allenby (2020) "Improving Text Analysis Using Sentence Conjunctions and Punctuation," *Marketing Science*, 39(4), pp.727–742.
- Carpenter, B., A. Gelman, M.D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li and A. Riddell (2017) "Stan: A probabilistic Programming Language," *Journal of Statistical Software*, 76(1).
- Chen, Y., R. Iyengar and G. Iyengar (2017) "Modeling Multimodal Continuous Heterogeneity in Conjoint Analysis — A Sparse Learning Approach," *Marketing Science*, 36(1), pp.140–156.
- Chintagunta, P., D.M. Hanssens and J.R. Hauser (2016) "Editorial—Marketing Science and Big Data," *Marketing Science*, 35(3), pp.341–342.
- Cui, D. and D. Curry (2005) "Prediction in Marketing Using the Support Vector Machine," *Marketing Science*, 24(4), pp.595–615.
- Davenport, T., A. Guha, D. Grewal and T. Bressgott (2020) "How Artificial Intelligence Will Change the Future of Marketing," *Journal of the Academy of Marketing Science*, 48(1), pp.24–42.
- Dew, R. and A. Ansari (2018) "Bayesian Nonparametric Customer Base Analysis with Model-Based Visualizations," *Marketing Science*, 73(2), 216–235.
- Dew, R., A. Ansari and Y. Li (2020) "Modeling Dynamic Heterogeneity Using Gaussian Processes," *Journal of Marketing Research*, 57(1), pp.55–77.
- Domingos, P. (2015) *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books: New York.
- Dzyabura, D. and J.R. Hauser (2011) "Active Machine Learning for Consideration Heuristics," *Marketing Science*, 30(5), pp.801–819.
- Everson, P.J. and E.T. Bradlow (2002) "Bayesian Inference for the Beta-Binomial Distribution via Polynomial Expansions," *Journal of Computational and Graphical Statistics*, 11(1), pp.202–207.
- Gabel, S., D. Guhl and D. Klapper (2019) "P2V-MAP: Mapping Market Structures for Large Retail Assortments," *Journal of Marketing Research*, 56(4), pp.557–580.
- Gilbride, T.J., G.M. Allenby and J.D. Brazell (2006) "Models of Heterogeneous Variable Selection," *Journal of Marketing Research*, 43(3), pp. 420–430.
- Halperin, I. (2017) "Inverse Reinforcement Learning for Marketing," SSRN: <http://dx.doi.org/10.2139/ssrn.3087057>.
- Hoffman, M.D., D.M. Blei, C. Wang and J. Paisley (2013) "Stochastic Variational Inference," *Journal of Machine Learning Research*, 14, pp.1303–1347.
- Huang, D. and L. Luo (2016) "Consumer Preference Elicitation of Complex Products Using Fuzzy Support Vector Machine Active Learning," *Marketing Science*, 35(3), pp.445–464.
- Ishigaki, T., N. Terui, T. Sato and G.M. Allenby (2018) "Personalized Market Response Analysis for a Wide Variety of Products from Sparse Transaction Data," *International Journal of Data Science and Analytics*, 5, pp.233–248.
- Jacobs, B.J.D., B. Donkers and D. Fok (2016) "Model-Based Purchase Predictions for Large Assortments," *Marketing Science*, 35(3), pp.389–404.
- Kucukelbir, A., R. Ranganath, A. Gelman and D.M. Blei (2015) "Automatic Variational Inference in Stan," *Neural Information Processing Systems*, pp.568–576.
- Lin, S., J. Zhang and J.R. Hauser (2015) "Learning from Experience, Simply," *Marketing Science*, 34(1), pp.1–19.
- Liu, J. and O. Toubia (2018) "A Semantic Approach for Estimating Consumer Content Preferences from Online Search Queries," *Marketing Science*, 37(6), pp.930–952.
- Miller, S.J., E.T. Bradlow and K. Dayaratna (2006) "Closed-Form Bayesian Inferences for the Logit Model via Polynomial Expansions," *Quantitative Marketing and Economics*, 4, pp.173–206.
- Molnar, C. (2019) *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*, <https://>

christophm.github.io/interpretable-ml-book/.

- Ng, A.Y. and S.J. Russell (2000) "Algorithms for Inverse Reinforcement Learning," in *Proceedings of the 17th International Conference on Machine Learning*, pp.663-670.
- Puranam, D., V. Narayan and V. Kadiyali (2017) "The Effect of Calorie Posting Regulation on Consumer Opinion: A Flexible Latent Dirichlet Allocation Model with Informative Priors," *Marketing Science*, 36 (5), pp.726-746.
- Rasmussen, C.E. and C.K. Williams (2006) *Gaussian Processes for Machine Learning*, The MIT Press.
- Ruiz, F.J.R., S. Athey and D.M. Blei (2020) "SHOPPER: A Probabilistic Model of Consumer Choice with Substitutes and Complements," *The Annals of Applied Statistics*, 14(1), pp.1-27.
- Rutz, O.J., M. Trusov and R.E. Bucklin (2011) "Modeling Indirect Effects of Paid Search Advertising: Which Keywords Lead to More Future Visits?" *Marketing Science*, 30(4), pp.646-665.
- Schwartz, E.M., E.T. Bradlow and P.S. Fader (2014) "Model Selection Using Database Characteristics: Developing a Classification Tree for Longitudinal Incidence Data," *Marketing Science*, 33(2), pp.188-205.
- Shiller, B.R. (2020) "Approximating Purchase Propensities and Reservation Prices from Broad Consumer Tracking," *International Economic Review*, 61(2), pp.847-870.
- Smith, A.N., P.E. Rossi and G.M. Allenby (2019) "Inference for Product Competition and Separable Demand," *Marketing Science*, 38(4), pp.690-710.
- Tibshirani, R. (1996) "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B*, 58(1), pp.267-288.
- 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.
- Wedel, M. and P.K. Kannan (2016) "Marketing Analytics for Data-Rich Environments," *Journal of Marketing*, 80(6), pp.97-121.
- Xia, F., R. Chatterjee and J.H. May (2019) "Using Conditional Restricted Boltzmann Machines to Model Complex Consumer Shopping Patterns," *Marketing Science*, 38(4), pp.711-727.
- Zhong, N. and D.A. Schweidel (2020) "Capturing Changes in Social Media Content: A Multiple Latent Change-point Topic Model," *Marketing Science*, 39(4), pp.669-848.
- Zou, H. and T. Hastie (2005) "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society. Series B*, 67(2), pp.301-320.
- 佐藤一誠 (2016) 『ノンパラメトリックベイズ：点過程と統計的機械学習の数理』 講談社
- 鈴木大慈 (2015) 『確率的最適化』 講談社
- 富岡亮太 (2015) 『スパース性に基づく機械学習』 講談社
- 平井有三 (2012) 『はじめてのパターン認識』 森北出版
- 持橋大地・大羽成征 (2019) 『ガウス過程と機械学習』 講談社