

博士論文

コンピュータービジョンと機械学習の統合による
モバイル広告効果の要因分析手法

A Method for Analyzing the Contribution of Mobile Advertising Effectiveness
through the Integration of Computer Vision and Machine Learning

2021 年度

指導教員 本橋 永至

(学籍番号 : d16vb002)

崎濱 栄治

Eiji Sakihama

横浜国立大学大学院
国際社会科学府 博士後期課程
経営学専攻

提出日 : 2021 年 12 月 1 日

目次

第 1 章	はじめに	1
1.1	背景	1
1.2	目的	6
1.3	構成	7
第 2 章	先行研究のレビュー	8
2.1	インターネット広告に関する研究	8
2.2	アンサンブル学習に関する研究	17
2.3	先行研究を受けて	23
第 3 章	コンピュータービジョンによる画像広告要素のクリック訴求効果の検証	25
3.1	はじめに	25
3.2	先行研究	27
3.3	使用データと分析対象	29
3.4	実証分析	33
3.5	本章のまとめ	39
第 4 章	アンサンブル学習と LDA の統合による動画広告の要因分析	41
4.1	はじめに	41
4.2	先行研究	43
4.3	使用データと分析対象	45
4.4	実証分析	48
4.5	本章のまとめ	53
第 5 章	おわりに	57
5.1	本研究のインプリケーション	57

5.2	本研究の貢献と限界	59
5.3	今後の展望	60

第 1 章

はじめに

1.1 背景

[Frey 13] の “The future of employment: How susceptible are jobs to computerisation?” は、人工知能 (AI) によって多くの職業が機械化され消滅するとの文脈で非常に話題を呼んだ (Google Scholar によると 2021 年 8 月 28 日現在の引用回数は 8,648 回)。同論文は、米国労働省 (United States Department of Labor) が作成している職種定義のオンラインデータベース、O*NET(Occupational Information Network)^{*1} を利用し機械化・コンピューター化がもたらす職種への影響を統計モデルによって評価している。結論のランキング表は人工知能 (AI) によって人がやるべき職業が奪われる・代替される…と引用されることが多いが、[Frey 13] ではあくまでも「機械学習 (Machine Learning:ML)」と「モバイルロボット (Mobile Robotics:MR)」による機械化・コンピューター化 (computerisation) による影響を検証するとしている点は注意が必要である。

同論文の手順は以下の通りである。特定の職種をコンピューターによって代替する際にボトルネックとなる 3 つのタスク (知覚と操作, 創造的知性, 社会的知性) を O*NET 内の変数と関連付け, 9 つの変数を特定する。具体的には以下通りである。

^{*1} 日本では、厚生労働省が日本版 O-NET として「職業情報提供サイト」(URL:<https://shigoto.mhlw.go.jp/User/>) を展開している。

1. 知覚と操作 (Perception and manipulation)

- 指先の器用さ (Finger dexterity)
- 手先の器用さ (Manual dexterity)
- 窮屈な作業スペース・不便な姿勢 (Cramped work space, awkward positions)

2. 創造的知性 (Creative intelligence)

- オリジナリティ (Originality)
- 芸術的な能力 (Fine arts)

3. 社会的知性 (Social intelligence)

- 社会的認知度 (Social perceptiveness)
- 交渉力 (Negotiation)
- 説得力 (Persuasion)
- 人を助けること・気遣うこと (Assisting and caring for others)

これらの変数は順序尺度によって定義されている。全 702 職種のうち 70 については主観的に機械化・コンピューター化による自動化が可能であれば 1 を、そうでなければ 0 を割り当てダミー変数とする。このデータを用いて分類器を構築し、残りの 632 職種について自動化可否の確率を推定している。表 1.1 に、自動化の確率が 99% とされた職種をあげている。

機械化・コンピューター化のボトルネックになるタスクを自動化することは困難を極めるとしても、テクノロジーによって人の支援を行うことは可能であろう。例えば、知覚と操作系タスクの窮屈な作業スペース・不便な姿勢については、例えば、サイバーデザイン社

表 1.1 機械化・コンピューター化される確率が 99% とされる職種

[Frey 13]

No.	職業名
1.	Data entry keyers
2.	Library technicians
3.	New accounts clerks
4.	Photographic process workers and processing machine operators
5.	Tax preparers
6.	Cargo and freight agents
7.	Watch repairers
8.	Insurance underwriters
9.	Mathematical technicians
10.	Sewers,hand
11.	Title examiners,abstractors,and searchers
12.	Telemarketers

の Web サイト^{*2}によると「HAL®腰タイプ作業支援用は、重量物を持ったときに、腰部にかかる負荷を低減することで、腰痛を引き起こすリスクを減らします。作業現場の労働環境改善や労働災害防止への活用が期待できます。」とする、同社の装着型サイボーグ製品が支援できるケースがあり得る。

創造的知性系タスクのオリジナリティ、芸術的な能力に関しては、人工知能学会誌が 2020 年 5 月に「AI でよみがえる手塚治虫」と題した特集の中で、[松原 20] の「「TEDUKA 2020」プロジェクト-人と AI が協力してマンガを描く-」、[手塚 20] の「AI は天才を生むか～人と AI の共同創作」といった報告がされており、人と AI が協力することによって新たな手塚作品を創造する挑戦がなされている。

最後の社会的知性系タスクについては、[石川 18] の「感情表現を用いた説得対話システ

^{*2} https://www.cyberdyne.jp/products/Lumbar_LaborSupport.html(accessed 2021/11/19)

ム」や、[稲葉 19]の「ユーザーに上手に説得される人狼知能対話システムの構築」などのように感情表現を伴った説得的な対話に関する研究が進捗している。どちらもデータセット構築においては、クラウドワーカーによる判断が必要なアノテーション作業が要求されている。

このように [Frey 13] で自動化のボトルネックとされたタスクの一部は、人と機械が協業を前提としながら実現に向けた改善・改良の取り組みがなされている。

[中島 18] は、国勢調査の分類を用いて消えた平成の職業について報告している。IT化の進展と深い消えた職業としては「ワードプロセッサ操作員、タイピスト」、「場立人、才取人」があげられる。「ワードプロセッサ操作員、タイピスト」はタイピングを専門とする職業であったが、PCの普及によりその技能が一般化したため専門職としての従事者が減少したこと、「場立人、才取人」は証券取引所の立会場で株式売買の仲介を行う専門とする職業であったが、1999年に株式売買が自動化されたため消滅した。また、近年の音声認識技術の発展に伴って「音声情報の聞き起こし」(通称:テープ起こし)の需要も減少していく可能性が高い。

他方、IT技術の進展によって平成で新たに誕生した職業として [中島 19] は、「システムコンサルタント」、「情報セキュリティ技術者」などのIT関連技術者や「金融商品開発者」、「金融ストラテジスト」などの金融系の職業をあげている。

今後、「場立人、才取人」のように、機械化・コンピューター化の進展に伴って消滅する職業がある一方で、IT関連技術者のようにそれまで存在していなかった職業が誕生することが想定される。人が高度なIT技術を使いこなし、創造力の源泉としたことで諸問題を抱えながらもスマートフォンや各種インターネットサービスが当たり前となる現代に到達したとも言えよう。

近年、目覚ましい発展を遂げている深層学習や機械学習関連のテクノロジーを応用する

ことで、存在価値が著しく低下する職業があるだろう。しかしながら、新しいテクノロジーを活用することで既存業務の効率化を支援する、人の創造力を拡張する、といった「人と機械との協業」の価値が高まっていくものと考えられる。

Brynjolfsson と McAfee は、「機械との競争 (原題:Race Against The Machine)」[Brynjolfsson 13] の中で人間と機械の理想的な協業事例としてチェスをあげている。チェスでは、人間がコンピューターと対戦しても勝つことができなくなったことから、人間とコンピューターを自由に組み合わせる「フリースタイル」が認められることとなった。世界一となったのは最強の人間でも最強のコンピューターでもなく、アマチュアの2名と3台のコンピューターからなるチームであった。アマチュアの2名はコンピューターに学習させる能力に長けており、「弱い人間+コンピューター+よりよいプロセス」の組み合わせが、「1台の強力なコンピューター」や「強い人間+コンピューター+お粗末なプロセス」に勝利したと報告している。

インターネット広告の分野においても、検索連動型広告 (リスティング広告) を対象に自然言語処理を活用して広告効果が高い (クリックされやすい、商品購入に繋がりやすいなど) とされるキーワードを特定することで、人間がより魅力的な広告文を効率よく作成できる支援方法について Google による事例が公開されている*3。

本研究では、インターネット広告における画像広告・動画広告を題材とし、テクノロジーを援用することで人の創造力を効率的かつ効果的に発揮できると期待される、現時点で実装可能な「人と機械との協業」方法を示す。

*3 機械学習のその先に初めの一步を踏み出そう:https://www.youtube.com/watch?v=jN_pRE0P1GU(access 2021/11/8)

1.2 目的

インターネット広告では、バナー広告・ディスプレイ広告と呼ばれる画像広告が一般的であったが、通信環境の改善、高性能なモバイル端末が広く普及したことを背景に、動画広告が利用されるようになってきている。[電通 20] の図 1.1 によると、2020 年の動画広告の市場規模は 3,862 億円 (前年比 21.3% 増) であり、検索連動型広告 (前年比 1.6% 増) とディスプレイ広告 (前年比 3.4% 増) と比較して大幅に成長している。動画広告は、1. 情報やイメージが印象に残りやすい、2. テキスト広告よりも多くの情報を伝えられる、3. 拡散力があり話題になりやすい、といった特徴を備えている。

本研究では、モバイル広告のデータを対象に、大手 IT 企業が提供するコンピュータービジョンサービス、機械学習の一つであるアンサンブル学習、ブラックボックスとも言わ

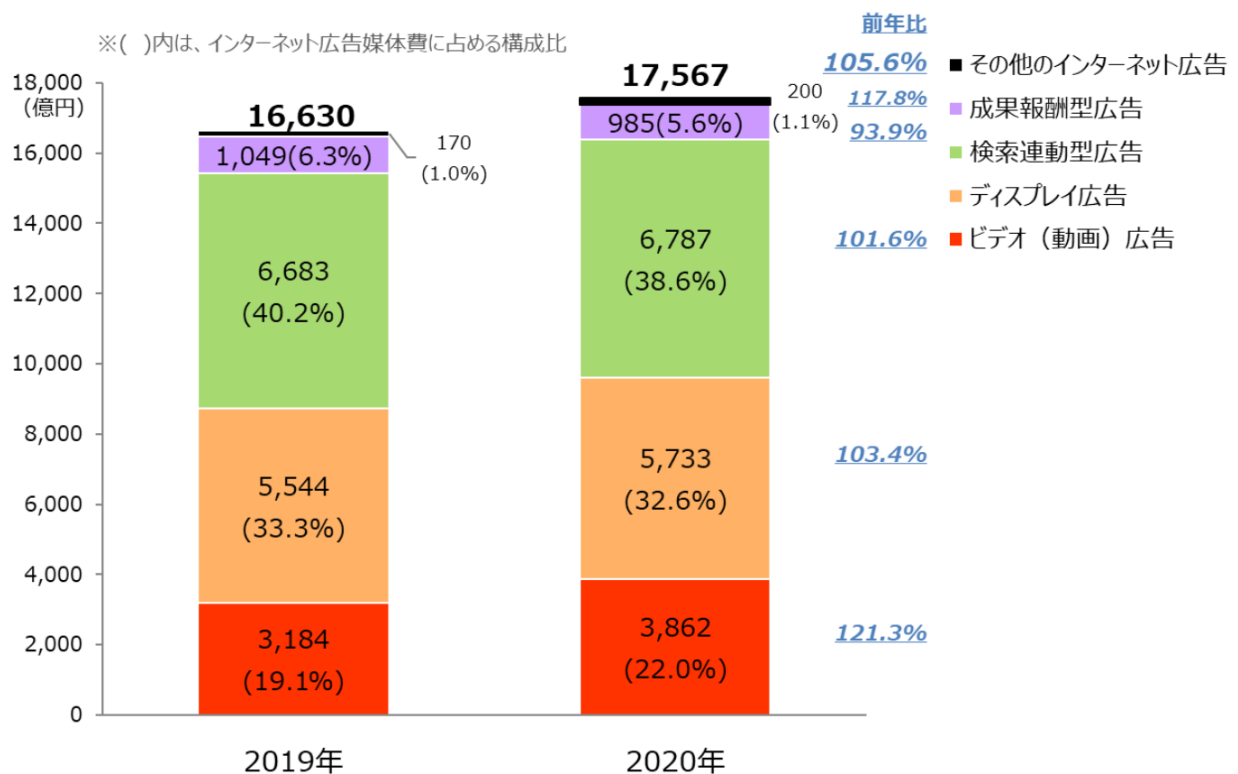


図 1.1 インターネット広告媒体費の広告種別構成比:[電通 20]

れる機械学習モデルの解釈を試みる説明可能 AI(Explainable AI:XAI), 自然言語処理を組み合わせることで, 効果的な広告クリエイティブ制作の支援を目的とする. 各特徴量の重要度だけでなく, 相互作用も含めルールベースでアンサンブル学習モデルの解釈を可能とする. また, 本研究の枠組みは, 広告効果に相当する教師データ (売上金額, 販売数量, 人による評価集計など) が得られるタスクであれば, 映像・画像・テキスト・音声を組み合わせて取り扱うことができることから, 販促用ポスター広告, 駅やオフィスビルにおけるデジタルサイネージ広告, タクシー内の動画広告など広告の他にも, TV コマーシャル, ドラマ・映画の予告編といった幅広い映像コンテンツの評価にも適用することが可能である.

1.3 構成

本研究の構成は以下の通りである, 第2章では国内外のインターネット広告とアンサンブル学習に関連する先行研究を取り上げる. 第3章では画像広告 (バナー広告, ディスプレイ広告) を対象とし, 第4章では動画広告を対象とした「人と機械の協業」手法について詳述する. 第5章では本研究のインプリケーション, 貢献と限界, 今後の展望についてまとめを行う.

第 2 章

先行研究のレビュー

本研究では、「人と機械の協業」可能性の例として、インターネット広告の画像広告・動画広告の素材(クリエイティブ)と広告配信に関するログデータを対象としていることから、まず、国内外のインターネット広告研究について概観する。続いて、本研究の分析モデルの軸となるアンサンブル学習の関連研究について概観する。アンサンブル学習とは複数の学習器の予測結果を組み合わせることで、単独の学習器の予測精度を上回ることを目指す機械学習手法の一つである。また、一般にブラックボックスと呼ばれる機械学習モデルの解釈を試みる研究分野に説明可能 AI(Explainable AI:XAI)がある。説明可能 AI 分野で対象とされる機械学習モデルとしては、学習器として決定木を用いたアンサンブル学習モデルと深層学習を用いたモデルが存在する。これらのモデルについて概観する。

2.1 インターネット広告に関する研究

2.1.1 はじめに

日本における 2020 年の総広告費は新型コロナウイルス感染症の影響を受け前年度比 88.8%と落ち込んだが、インターネット広告は前年度比 105.9%と、新聞やテレビメディア広告が減少傾向にある中で唯一の増加傾向にある [電通 20]。これは日本だけではなく

海外も同様の傾向であることから [電通 21], 本節では国内外で成長が続くインターネット広告について, 定義や大まかな分類を整理した上で, これまでなされてきた研究を時系列に沿って示す.

2.1.2 インターネット広告の定義

インターネット広告の定義は, 一般的には富士通総研が示すように「インターネット広告とは, 基本的に Web サイト上や電子メール内のスペースに掲載される広告のこと」[富士通総研 21] であるが, 自社の HP にて製品・サービスをプロモーションすることや, 自社のメールマガジンを広義のインターネット広告とする捉え方もあり, その定義は曖昧である. 例えば [戸田 10a] も歴史的変遷を整理する前提でインターネット広告について, 「広告の範囲と定義が流動的であり曖昧なまま」と述べている. 本節では, 基本的には富士通総研の定義に即すと共に, 自社 HP でのプロモーションやメールマガジンは定義から除くものとする.

2.1.3 インターネット広告の種類

インターネット広告の種類や分類は多岐にわたる状況であるため, ここでは代表的なイメージ (図 2.1) とその概要を示す.

Web サイト上や動画再生時, アプリ実行時に表示される広告のことをディスプレイ広告と呼び, その多くがバナーと呼ばれる画像形式であるためバナー広告とも呼ぶ. Web サイトやアプリに掲載される場合は画像, 動画といった形式や, コンテンツ間に配置するインフィールド広告といった表示形式など, 様々な種類が存在する. また動画サイトの場合は, 動画再生前又は途中で表示される広告もディスプレイ広告の一種である.

ディスプレイ広告は, インターネット広告の初期からある広告形式であり, 現在も広く使用されているが, 図 1.1 で確認した通り近年は検索連動型広告のシェアがディスプレイ



図 2.1 ディスプレイ広告と検索連動型広告のイメージ:[図書印刷 20]

広告を上回っている。ディスプレイ広告は基本的には、従来の屋外広告やテレビ広告と似た利用者のアクセスを待つ受動型の性質を有することから、ニーズが固まっていない潜在層の利用者向けに使用されることが多い。

一方で検索連動型広告とは、利用者が検索エンジンに入力したキーワードに連動した広告が表示される形式である。テキスト形式がほとんどであり、ディスプレイ広告が Web サイトに表示される一方、検索連動型広告は検索結果画面の上部に表示される。このようにニーズが明確な利用者を対象とする形式のため、ニーズの顕在層向けに使用されることが多く、また顕在層にアプローチするため有効性が高いとされている。なお、検索連動型広告をリスティング広告と言い換えることもあれば、ディスプレイ広告と検索連動型広告の双方を含めてリスティング広告と呼ぶこともある。

2.1.4 インターネット広告とその研究の変遷 (1990 年代後半～2000 年代前半)

[戸田 10a] によればインターネット広告が日本に登場したのは 1996 年からである。1990 年代はインターネットそのものも黎明期であり、ディスプレイ広告もインターネットの進歩に歩みを合わせつつ、様々な手法が試みられた時期であった。そのためインターネット広告に関する研究は、国内外で基礎的な研究が中心であった。

インターネット広告の黎明期の中で、1990 年代後半から 2000 年代前半にかけての日本

では、[中畝 98]によるインターネット広告の効果測定に向けた指標の検討や[形山 01]による広告手法の日米比較、[井上 02]による有効性の検証などといった基礎的研究がなされた。[中根 04]はインターネット広告の現状や手法、課題について、市場調査やアンケート調査を通じて整理した上で今後の展望を示唆しており、検索連動型広告の可能性と発展性について言及している。

海外においては、インターネット広告の登場は日本よりも若干早いものの、基礎的研究が中心であるという傾向は同様であった。例えば[Schlosser 00]は利用者のインターネット広告に対する基本的な評価・態度の検証をし、[Previte 98]はインターネット経験年数がインターネット広告への評価に影響が及ぼすことを立証した。一方、日本の研究傾向と異なる点としては、2000年代前半においてインターネット広告の回避に関する研究が先駆けて行われていたことがある。[Cho 04]は利用者がインターネット広告を回避する3つの要因として、「知覚された目標阻害」、「知覚された広告の散乱」、「過去のネガティブな経験」をあげた。

2.1.5 インターネット広告とその研究の変遷(2000年代後半～2010年代前半)

2000年代の大きなトピックは、2002年にGoogleが開始した検索連動型広告のサービス、及び2006年に日本に導入された利用者の行動履歴をもとに広告を表示する行動ターゲティング広告である。これらのサービスの登場によって、インターネット広告の市場におけるシェアがさらに増加していった。

ただし研究においては、検索連動型広告などに関する有効性の検証はさして行われていない。ニーズが顕在化している利用者への広告であるため、その有効性は自明であったからであろう。

検索連動型広告に限定しなければ、インターネット広告の有効性に関する研究は継続的に行われており、[中岡 08]はディスプレイ広告に着目して感性評価実験及び因子分析

を行い、クリックしたくなる広告の要因抽出を試みた。また、[堀田 08] は GPS 機能と連動した新たなインターネット広告の仕組みとその有効性を検証した。[岡田 12] は、バナー広告のデザインに有効な感性データをクリック率から特定する実験を行った。なお、1990 年代後半～2000 年代前半と比べると、インターネット広告全般ではなく特定の広告手法を対象とした研究や、新技術の効果検証を行う研究が増加したと言える。

海外においては、検索連動型広告に特化した論文は国内よりも比較的多い。例えば、[Ghose 09] は連立方程式モデルなどを用いてクリック率やコンバージョン率、広告の掲載順位などといった複合的な関係を分析した。[Animesh 11] は、リスティング広告のクリック率に関するモデルを構築し実際のデータから分析したことで、ライバル企業との差別化の程度などがモデルに影響を与えていることを明らかにした。

[Mcmahan 13] は Google の検索連動型広告 (リスティング広告) におけるクリック率 (Click Through Rate:CTR) 予測の研究において、データが追加される毎に逐次学習を実行する Online Gradient Boosting(オンライン勾配ブースティング) の発展形を提案した。また検索連動型広告に限らず、ディスプレイ広告に関する研究も継続して行われている。[He 14] は、Facebook 広告を対象として Gradient Boosting Decision Trees:GBDT(勾配ブースティング決定木) とロジスティック回帰を組み合わせたクリック率予測手法を検証している。[Nihel 13] はインターネット上の情報量が莫大であるという前提のもと、広告の効率性を検証するため、「ディスプレイ広告のクリック」と「利用者の記憶」に影響を与える要因は画像の色、位置、サイズ、メッセージの明確さによることを実験により明らかにした。

この時期から効率化に向けた研究が行われ始め、例えば [山川 11] は画像広告を要素に分解して要素別に様々なパターンを用意した上で、A/B テストによる検証を行った。[田頭 13] は、テキスト広告 (広告自体がテキストのみからなる広告) を対象に特徴量のグルー

個別に比較検討を行った。[宮西 14] は、複数の機械学習手法を組み合わせることで、ユーザーの興味・関心と広告属性の両面から適切な広告を推薦する手法について実証実験を行い有効性を確認した。[馬場 15] によって広告に使用するテキストの自動化及びその評価指標の検討がなされた。人工知能 (AI) などを用いたモデルによる効率化の傾向は 2010 年代後半の大きなトピックであるため、次節においても詳述する。

インターネット広告の回避や利用者の保護といった観点での研究が日本で現れ始めたのもこの時期である。[西村 10] は検索連動型広告の有効性を整理した上で、利用者へのアンケート調査を実施し検索連動型広告を回避する場合の要因分析を行った。[近江 15] はインターネット広告の発展性を認識した上で、法律に抵触する恐れがある内容などを有害広告と位置づけ、その対応策としてオプトアウト方式の可能性を提案した。海外においては、利用者の保護についても継続して研究が行われており、[Tian 15] は、特にクラウドソーシングにおける詐欺被害を引き起こす不正広告の実態を整理した上で、不正広告の検出システムの構築及びその有効性を検証した。

[戸田 10a,10b] が 1990 年代から 2010 年時点におけるインターネット広告の変遷を詳細に整理しており、国内外の概況を把握することができる。その他、変遷に関する研究としては [Jansen 08] が検索連動型広告に特に着目し、その歴史と技術の変遷を概観した上で、クリック詐欺やオークション形式の問題点、今後の展望について整理した。

2.1.6 インターネット広告とその研究の変遷 (2010 年代後半～)

2010 年代の後半はインターネット広告の手法としては大きな変化は無いものの、依然として広告費に占めるインターネット広告の割合は増加しており、その役割は大きくなるばかりである。

そのため日本においては、基礎的研究よりも AI などを用いた効率化を図る研究が活発化していると言えるだろう。すなわちインターネット広告の市場規模は未だ拡大傾向に

あり、いかに省力化し効率的な運用をするかということが課題となっていると思われる。また合わせて、日本においても広告回避や利用者保護といった観点の研究が充実してきていることも研究傾向の特徴である。一方、海外においても効率化を図る動きは同様であるが、インターネット広告そのものの信頼性について問う研究が増加したことは特徴的であろう。

2017年には「広告とAI」と題した特集が「人工知能学会誌」にて生まれ、広告活動の自動化や品質向上に資するAIの役割について、各種の取組が紹介された。特に本特集の中で[谷口 17]はインフィード広告内における画像の分析と自動生成、及び動画広告の自動生成に関する技術について詳述している。

このようにAIや深層学習に対する広告業界の期待値は大きく、他にもFacebookでの広告を題材にクリック率予測の実験を行った[岩崎 18]の研究や、特定のキーワードを抽出した上で広告文を自動生成する技術を検証した[脇本 20]の研究などがあげられる。また、[Narita 19]は広告配信ログデータを利用して、広告配信実験をせずに新しい広告クリエイティブを評価する手法を提案している。海外に目を向けると、一般的な広告配信ログといった特徴量に加えて、画像広告や動画広告のピクセル情報やテキストについてもDeep Learningによって統合的に取り入れた研究に[Chen 16a]や[Lou 18]がある。

利用者の保護に関して、[若江 19]はトラッキングに伴う個人情報流通の現状について改めて整理した上で、現行法に基づく規制の可能性及び法改正の必要性について言及した。このように法的観点からインターネット広告を論じた研究が登場した点は、2010年代後半の特徴と言える。

加えて、[李 21]はユーザーがインターネット広告を回避する傾向について国内外の研究を体系的に整理した上で、Web調査に基づき回避を規定する要因を調査し、広告手法や表現による不快感及びインターネット利用を妨害するインターネット広告の特徴が回

避を高める重要な要因であることを明らかにした。

インターネット広告の有害性や上述した回避的態度とは対比的に、インターネット広告を商用目的以外に、かつ他分野で活用するといった観点で、自殺関連語を検索する利用者への援助要請行動を促しやすいインターネット広告のあり方を検討した [高橋 20] の研究は特徴的である。 [Crain 19] は、インターネット広告が諸外国から干渉・悪用され選挙に影響を与えていることを前提に、インターネット広告含むデジタルインフラのあり方を論じている。また [Borisova 17] は、インターネット広告と屋外広告とを改めて比較し、インターネット広告は利用者から人気があるものの信頼性が低く、場合によっては屋外広告に注力した方が良いことを明らかにした。このようにインターネット広告の影響力を前提としながら、その信頼性に着目した研究についても実施されている。

2.1.7 本節のまとめ

本節では3つの時代区分を設け、それぞれにおける研究の傾向を概観した。以下にその傾向の概要を改めて示す。

- 1990年代後半～2000年代前半
 - ・国内外共にインターネット広告の効果などを測定する基礎的な研究が主である。
 - ・海外ではインターネット広告への回避行動に関する研究が先行して行われた。
- 2000年代後半～2010年代前半
 - ・2000年代前半の検索連動型広告の登場に伴い、国内外でインターネット広告を効率化するモデル研究が活発化した。
 - ・国内においてもインターネット広告の回避行動に関する研究が実施され始めた。
- 2010年代後半～
 - ・AIなどを用いたインターネット広告の効率化が国内外でさらに充実した。

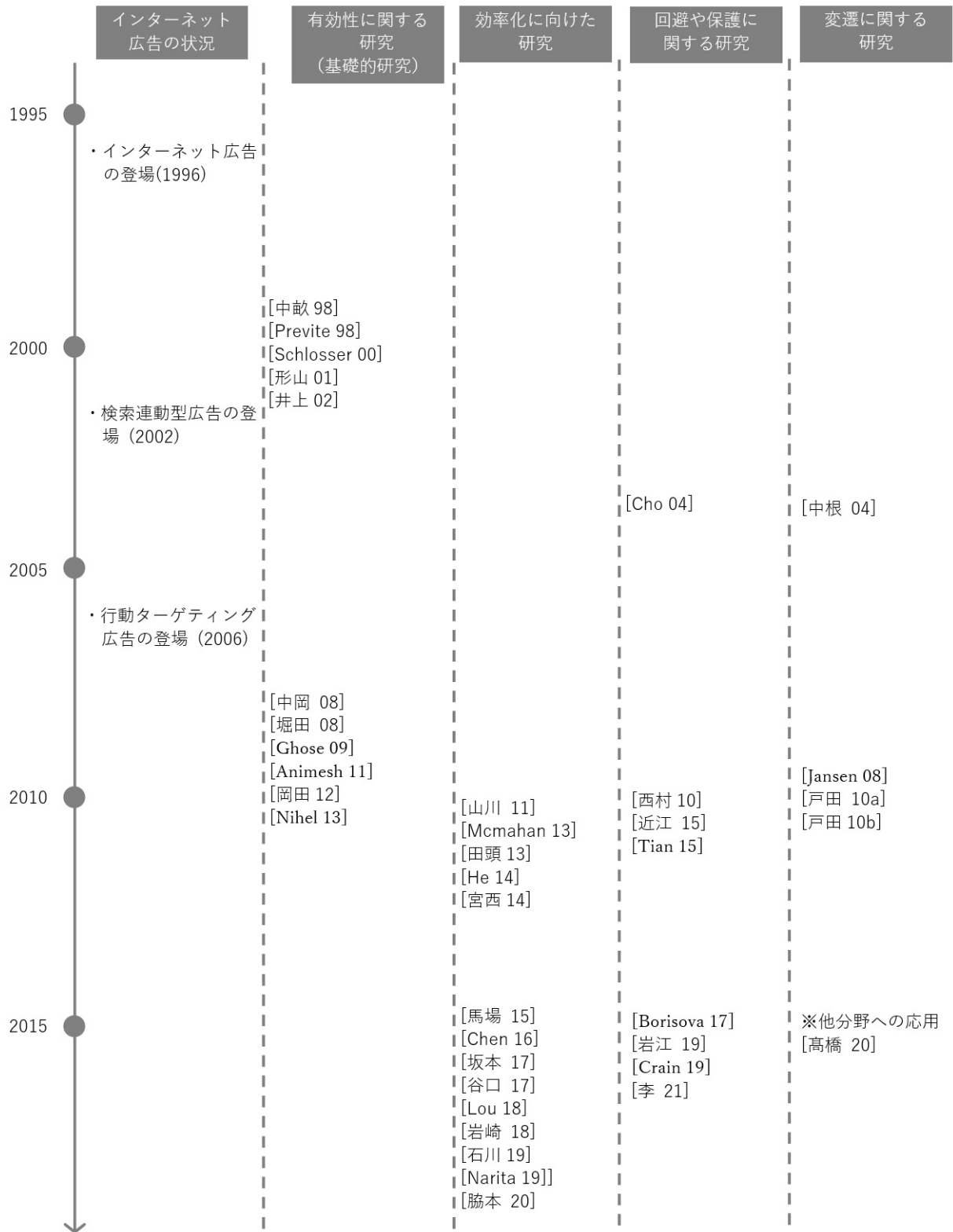


図 2.2 インターネット広告及びその研究の分類と変遷

・海外においては、インターネット広告の信頼性について改めて検証する動きがある。

2.2 アンサンブル学習に関する研究

2.2.1 アンサンブル学習の特徴

アンサンブル学習とは、複数の学習器 (弱学習器, Weak Learner) を組み合わせることで意思決定を行う手法の総称であり、主に教師あり学習のタスクにおいて用いられる。決定木、ニューラルネットワーク、線形回帰、ロジスティック回帰など様々なモデルが学習器として用いられ、複数の学習器を組み合わせることで単独の学習器による予測精度を上回ることが期待される。アンサンブル学習は、「群衆の知恵」を機械学習に置き換えたものであると言える [Sagi 18]。

[Surowiecki 04] によると以下の4条件を満たす場合、群衆の知恵は一人の意思決定者を凌駕する可能性が高いとされる。

- Independence(独立): One's opinion is not affected by other opinions.
- Decentralization(分権): One is capable of specializing and making conclusions based on local information.
- Diversity of opinions(意見の多様性): One should hold private information, even if it is just an eccentric interpretation of the known facts.
- Aggregation(集約): Some mechanism exists for turning private judgments into a collective decision.

アンサンブル学習が有効な理由について [Dietterich 02, Polikar 06] は、以下の点をあげ

ている。

- Overfitting avoidance(オーバーフィッティングの回避): 学習器を組み合わせることで誤った結果を選択するリスクを減少させ、全体の予測精度を向上させることができる。
- Computational advantage(計算上の利点): 局所的な探索を行う単一の学習器は局所的な最適値に陥る可能性があるが、アンサンブル学習では、複数の学習器を組み合わせることで局所解に陥るリスクを減らすことができる。
- Representation(表現): 最適な学習器は、単独の学習器の空間外に存在する可能性がある。異なるモデルを組み合わせることで探索空間を拡張することができる。その結果として、データ空間への適合性が向上する。

続いて、アンサンブル学習における代表的な手法について述べる。

2.2.2 アンサンブル手法

アンサンブル手法は、学習器の出力結果が次の学習器の学習時に影響を与える「依存型フレームワーク」と各学習器が独立して学習する「独立型フレームワーク」の2つに大別される。

[Freund 95] による AadaBoost は「依存型フレームワーク」として広く知られたアンサンブル手法の一つである。AdaBoost の特徴はひとつ前の学習器で誤って分類されたデータの重みを調整してから次の学習器による学習を実行することで、予測精度の向上を企図した手法である。ただし、AdaBoost は異常値やノイズの影響を受けやすい特徴がある。この点を考慮した発展形がその後の研究成果から存在しており、例えば、Soft margin AdaBoost[Rätsch 01] は、外れ値の影響を緩和するためにアルゴリズムに正則化を用いている。

[Breiman 96] による Bagging(bootstrap aggregating) は、「独立型フレームワーク」によるアンサンブル手法である。ブートストラップ法によって生成された学習データを用いて多数の学習器による学習を行い、教師データがカテゴリカルデータの場合は分類木による多数決を予測結果とし、連続データの場合は回帰木による平均値を予測結果とする。バギングアルゴリズムについても多数のバリエーションが存在し、[Jiang 11] による改良型バギングアルゴリズム (Improved Bagging Algorithm:IBA) がある。IBA では、各サンプルに情報エントロピーを付与することで、リサンプリング処理を改善している。

Random Forest は、[Amit 94] と [Ho 95] によってほぼ同時期に発表され、[Breiman 01] によって精緻化された「独立型フレームワーク」によるアンサンブル手法である。Random Forest はアンサンブル手法の中でも特に広く普及しており、例えば [Breiman 01] は、Google Scholar によると 82,000 回近く引用されている (2021 年 11 月現在)。Bagging との違いは、学習器による学習時の特徴量 (説明変数) 選択をランダムに実施する、各学習器の決定木の深さ (分岐) に制限を設ける、の 2 点である。結果として、似たような学習器が得られにくくなり、過学習による汎化性能の低下を防ぐことが期待される。

[Friedman 01,02] による Gradient Boosting Decision Tree:GBDT(勾配ブースティング) は、分類木で誤分類されたデータや回帰木で誤差の大きかったデータを改善 (設定した損失関数の値が小さくなる) する方向で、新たなモデルが構築される特徴を持つ「依存型フレームワーク」のアンサンブル手法である。[Chen 16b,Zhang 18] による XgBoost(eXtreme Gradient Boosting) は GBDT の発展形であり、評価関数を工夫することで過学習を低減し更に計算速度の向上を図っている。

本研究では、予測精度の高さと次節で述べる「説明可能 AI」の実装について考慮し、第 3 章では GBDT を、第 4 章は XgBoost によるモデル構築を行う。

2.2.3 アンサンブル学習のインターネット広告研究における適用例

Google の検索連動型広告 (リスティング広告) におけるクリック率 (Click Through Rate:CTR) 予測の研究において、データが追加される毎に逐次学習を実行する Online Gradient Boosting(オンライン勾配ブースティング) の発展形が提案されている [Mcmanhan 13]. [Çakmak 19] は、オンライン旅行予約サイト (Online Travel Agent:OTA) から各ホテル広告の表示場所、宿泊料金が最低価格となった回数、ホテルのレーティング情報など複数の特徴量を用いてクリック率予測の検証を行った。予測モデルの構築にあたって、Random Forest, GBDT, AdaBoost, Support Vector Regression, XgBoost の5つの性能比較を行い、XgBoost が最良であることが確認された。[Moneera 21] は、Kaggle(世界最大の機械学習コンペティションプラットフォーム) に提供されたインターネット広告の “Click-Through rate prediction dataset.” を対象に、K Nearest Neighbor, Logistic Regression, Random Forest, XgBoost による性能比較を実施した。実験結果から、XgBoost は特徴量の数を減らしても良好な予測性能があるとした。

2.2.4 説明可能 AI(Explainable AI:XAI)

本研究では、機械学習モデルの結果に加えて、その結果に関連する追加情報 (モデルの解釈、判断根拠の説明など) を対象とする研究分野全体を指す用語として説明可能 AI を用いる。なお、XAI は米国国防高等研究計画局 (Defense Advanced Research Projects Agency:DARPA) が主導している研究プロジェクトにおいて使われている略称である。同分野では国内においても [大坪 21] や [森下 21] が相次いで発刊されるなど注目を集めている。

[原 18,19] によると説明可能 AI の研究は、以下の4つに分類される。

1. 大域的な説明

2. 局所的な説明

3. 説明可能なモデルの設計

4. 深層学習モデルの説明

大域的な説明とは、複雑なブラックボックスと例えられるモデルを、人にとって可読性の高い解釈可能なモデルで表現することで説明とする方法を指す。深層学習モデルやランダムフォレストのような弱学習器として決定木を活用したアンサンブル学習が対象となる、例えば、単一の決定木やアソシエーションルールで近似的に表現することでモデルの説明とする。相互作用を考慮したルール抽出が可能であり画像広告や動画広告における広告効果検証において有益であると考え。本研究では、弱学習器として決定木を用いたアンサンブル学習モデルを用い、大域的な説明にあげる [Hara 18] と [Deng 19] によるルールベースによる解釈を行う。インターネット広告研究において、これらの手法を用いた特徴量の評価は調査した限りでは確認できていないことから新規性は高いものと考え。

Born Again Trees は、ニューラルネットワークなどのブラックボックスモデルを用いて追加の教師データを生成し、生成した新規データによって決定木の学習を行う [Breiman 96]。Node Harvest は、ランダムフォレストを分岐が浅い決定木で近似することで説明とする [Meinshausen 10]。[Hara 18] による、Making Tree Ensembles Interpretable-A Bayesian Model Selection Approach は、確率的なモデルとみなしたランダムフォレストを、単純なルールモデルへと変換する。[Deng 19] の Interpreting Tree Ensembles with inTrees は、ランダムフォレスト内の頻出ルールを主要ルールとして抽出し、ランダムフォレストの近似的な説明とする。なお、arXiv への投稿は 2014 年であり [Hara 18] 以前に提案されている手法である。

局所的な説明とは、ある特定の入力に対するモデルの予測根拠を示すことで説明とす

る方法を指す。ある入力 x をモデルが y と予測したときに、その予測の根拠を説明として提示する手法である。[Ribeiro 16] は、LIME(Local Interpretable Model-agnostic Explanations) と呼ばれる、線形モデル/ルールモデルを用いた局所的な説明を生成する方法を提案した。[Lundberg 18] は、SHAP(SHapley Additive exPlanations) と呼ばれる、いくつかの局所的な説明法 (例:LIME) がゲーム理論分野の重要な概念である Shapley Value の枠組みを利用して記述できることを示した。[Koh 17] は、予測に対して密接な関連を持つ訓練データを予測根拠として提示する方法。効率的な計算法を提案し、[Tolomei 17] は、LIME などのように「判断根拠になった特徴量」を説明するのではなく、判断結果を反転させる「“有用” な特徴量」を出力する方法を示した。また、[Lage 18] は局所的な説明を“わかりやすい”モデルの学習に応用する方法を提案した。他には、LIME では説明対象データの周りで線形モデルを作るが、Anchor では領域モデル (決定木の葉ノードのようなもの) による説明を試みた手法がある [Ribeiro 18]。

そもそも最初から可読性の高い解釈可能なモデルを作ってしまう方法として説明可能なモデル設計がある。上述の大域的な説明、局所的な説明、これら2つのアプローチはモデルを対象として説明を抽出することを目的としている。これに対し、この第3のアプローチ方法では初めから可読性の高い解釈可能なモデルを作ることを目的とする。[Angelino 17] や [Lakkaraju 16] はルールリストと呼ばれる決定木の派生モデルの学習方法を提案し、[Bien 11] は分類問題の各カテゴリの代表的な訓練データを特定する方法を示した。また、[Kim 16] は各カテゴリを代表する訓練データだけでなく、例外的な訓練データも合わせて提示する方法を提案している。

深層学習モデルの説明は、アプローチとしては2の局所的な説明に該当する。深層学習モデルの説明は、特に画像認識の分野で数多く研究されている。モデルが画像内のどの部分を認識しているかを特定してハイライトすることで説明としており、人の目によ

る確認を前提としている [Springenberg 14, Bach 15, Chen 16b, Sundararajan 17, Smilkov 17, Shrikumar 17].

2.3 先行研究を受けて

上述のように、インターネット広告の効率化に関する研究において、[谷口 17] は画像情報、動画情報、文字情報などを統合して広告効果との関連を分析した上で自動生成技術に詳述している。「人と機械の協業」に関連する研究として、[岩崎 18] の広告効果と画像広告領域との関連を可視化する研究をあげた。しかしながら、それらの結果から示唆を得るためには、広告クリエイティブを制作する際にクリエイターが一つ一つの画像広告を目視で判断する必要がある。また、画像広告や動画広告において描かれている物体と広告効果との関連を調査するために、人が一つ一つ目視で特徴量を付与することはコストや時間の観点から困難を伴う。よって、画像広告や動画広告の特徴量をコンピュータビジョンによって自動的に付与し、広告効果と特徴量との関係を人が理解しやすいキーワードとして抽出することができれば、クリエイターがより効率的に広告クリエイティブの制作を実行できる。

コンピュータビジョンによって動画素材から特徴量を抽出し予測モデルを構築した研究に [Kaminski 20] があげられる。同研究では、クラウドファンディングサイトを対象としてコンピュータビジョンを用いて動画素材から特徴量を抽出し、クラウドファンディングの成否をロジスティック回帰によって予測モデルを構築することで特徴量の重要度を示した。ただし、特徴量間の交互作用については今後の課題とされた。

本研究は、画像広告と動画広告を対象として、コンピュータビジョンによる特徴量抽出、弱学習器として決定木を用いたアンサンブル学習、大域的な説明モデル [Hara 18] と [Deng 19] による交互作用を考慮した特徴量の広告効果の可視化を行う。これらの結果が

ら，人が容易にモデルの解釈を行い効率的な広告クリエイティブ制作が可能となると期待される。

第3章

コンピュータービジョンによる画像広告要素のクリック訴求効果の検証

3.1 はじめに

インターネット広告は日常生活に浸透しており、リスティング広告、ディスプレイ広告、各種SNS広告、動画広告など日々進化を遂げている。その中で、Real Time Bidding(RTB)と呼ばれる広告表示の都度、オークションがリアルタイムで決定される仕組みは、広く普及してきた。適切な人に/適切な広告を/適切なタイミングで配信することを意図して、クリック率(CTR)の予測結果から入札価格を決定し、入札を経て広告が配信される。CTR予測に関する既存研究では、特徴量の追加や加工方法の検討、深層学習を含めた新たな機械学習手法の提案など膨大な研究成果が報告されている。近年は、広告配信のログデータに加えて、広告のテキストや画像情報についても包括的に特徴量として扱う研究画像が進められている[岩崎 18]。画像広告と配信に関する研究では、画像広告を要素に分解して、要素別に様々なパターンを用意した上で、A/Bテストや多腕バンディットアルゴリズムによる最適な組み合わせを提案する方法がある[山川 11, Narita 19]。

一方、画像広告のクリエイティブ制作はデザイナー個々人の勘と経験に依存する点も多く、実際に広告を配信してみなければ検証し難いという問題がある。デザイナーがクリ

エイティブ制作時にあらかじめ指針となる特徴量を理解していれば、必要な要素を意図的に要素として盛り込むことが可能となる。更に、新規に製作した画像広告について、意図的に盛り込んだ要素が客観的に反映できているか確認することもできる。

本章の目的は、画像広告から人が解釈可能なキーワードや色彩情報を抽出し、広告効果(クリック)に対して重要な要素を特定する手法の提案である。画像広告から情報を抽出する際には、コンピュータービジョンを利用する。深層学習研究の進展を背景に、[Schroff 15]では同年3月に画像認識で人間の精度を超える研究結果を報告した。画像認識の発展の背景としては、深層学習関連の研究成果が汎用的なライブラリとして組み込まれ、画像や動画について分析することが比較的容易になったことがあげられる。本章で利用する Microsoft Computer Vision API は、画像を分析し、人間が解釈できる言葉で、信頼度スコアと共に複数の情報が提供される。Amazon, Google, Microsoft といった外資大手 IT 企業は保有するビッグデータを活かして、API^{*1}サービスを展開している。Microsoft Computer Vision(MCV)の Web サイト^{*2}にアクセスし、画像を送信するとどのような要素が含まれているか、人間であれば性別や年齢、支配的な色調、画像の大きさといった分析結果を瞬時に確認することができる。これらサービスを利用するメリットは、自ら画像や動画などの教師データを大量に用意し、学習モデルを構築する必要がないことがあげられる。

本章の構成は以下のとおりである。次節では、CTR 予測問題、画像広告評価、コンピュータービジョンの応用の観点から先行研究のレビューを行う。第3節では、使用した広告配信ログと画像広告データに関して説明する。第4節では、訓練データの加工方法と検証に利用する学習機について、AUC(Area Under Curve)による予測精度比較を行

*1 Application Programming Interface の略:広義の意味ではソフトウェアコンポーネントが互いにやりとりするのに使用するインタフェースの仕様

*2 [https://azure.microsoft.com/ja-jp/services/cognitive-services/computer-vision/\(2017/7/17\)](https://azure.microsoft.com/ja-jp/services/cognitive-services/computer-vision/(2017/7/17))

う。次に、得られた学習機とテストデータの検証結果から、CTR 予測に対して重要な画像広告の構成要素とその相互作用について特定する。最後に、第 5 節において本章の課題を示す。

3.2 先行研究

CTR 予測問題ではログデータの性質上、千～万単位の特徴量と億単位のレコード数を扱うことも頻繁に行われる。様々な広告配信手法について A/B テストなど配信実験による検証が幅広く行われ、予測精度の向上を通じたビジネス貢献が目的とされている。[McMahan 13] では、「CTR 予測は数十億ドル規模のオンライン広告業界の中核をなす、大規模な学習問題である」と、その重要性をあげている。同論文では、Google のリスティング広告データを対象とし、CTR 予測における学習パラメーターの更新アルゴリズムについて、一般的な OGD(Online Gradient Descent:オンライン確率降下法) に代わって、FTRL-Proximal(Follow The Regularized Leader- Proximal) を提案した。[田頭 13] は、Yahoo! のディスプレイアドネットワークのテキスト広告を対象に、特徴量のグループ別に AUC による比較検討を行った。[山川 11] は、インターネット広告における画像要素の組み合わせ事例手法について報告した。更に、一般的な広告配信ログといった特徴量に加えて、画像広告のピクセル情報やテキストについても Deep Learning によって統合的に取り入れる研究が進められている [Chen 16b]。[岩崎 18] は、Facebook 広告を対象に、配信ログデータに加えて画像やテキスト情報を特徴量とした深層学習による CTR 予測を提案した。

画像広告そのものの評価という観点からは、アンケート調査や実験に基づく研究が行われている。[中岡 08] らは、アンケートによりバナー広告について 10 組の形容詞による感性評価を行った後、因子分析の結果から「目にとまる」「心をひかれる」の 2 因子を抽

出し、感性評価とクリックとの関連を分析した。[岡田 12] は、バナー広告のデザインに有効な感性データを得るために「オーバーラップ時間」「ワイプのパターンと方向」「有彩色と無彩色の配色効果」について実験を行いクリック率への影響を検証した。[石井 16] は、515 種類のバナー画像広告を収集し印象を「眼を惹く」「洗練されている」「分かりやすい」の 3 項目からアンケート調査を行い、その画像特徴量から判別する方法を提案している。[Wu 16] らは、バナー画像広告の視覚的な複雑さと他の画像広告との対比について実験による検証を行い、情報が詰め込まれた画像広告よりも目立つ方がユーザーに好まれるとの結論を得た。デザイナーの審美眼を明示的に扱った研究として、[石川 19] がある。Convolutional Neural Network(CNN)を利用して画像、色彩、文字に関する特徴量を抽出し、デザイナーが事前に判定した画像広告ラベル(良/悪)を教師データとする手法を提案し、特徴量重要度を報告している。同手法では、新規の画像広告の良/悪を判定できる可能性がある。一方、重要度が高い画像ベクトルの番号を解釈することは困難である。実際の広告配信ログと画像広告の構成要素について、人が解釈可能な文字情報の観点から分析された既存研究は見当たらなかった。

コンピュータービジョンの適用分野としては、画像分類があげられる。[Richards 18] は Google Cloud Vision(GCV)の結果と階層クラスタリングを用いて 20,000 枚の写真のグルーピングを行い、170 時間かかる手作業を効率化できるとした。[D'andrea 19] は、GCV による画像内容に則した検索手法を提案した。2018 FIFA World Cup Final Draw ceremony の Twitter に投稿された画像について分析を行ない、言語や地域の壁を越えられる可能性を示唆した。GCV と MCV を活用した研究分野に障害者支援がある。[Sivan 16] は、これまでの障害者支援技術は特定の領域に特化した複雑なデザインで商用としては高額であったとし、コンピュータービジョンを活用した安価な支援デバイスの提案を行なった。[Reis 18] は、GCV と MCV を併用した視覚障害者の生活品質向上に役立つ、

コンピュータービジョン技術の応用例を提案した。具体的には、1. 装着した Smartglass と呼ばれるメガネ型のデバイスから GCV と MCV に画像を送信、2. 画像の分析結果をスマートフォンで受信、3. Google 翻訳で英語からポルトガル語に変換、4. 変換したテキストを読み上げ機能で視覚障害者に情報を伝えるものである。

3.3 使用データと分析対象

3.3.1 使用データ

2019年6月24日から6月30日までの、1週間分のスマートフォン向けのモバイル広告配信ログを元に、Impression(表示回数)が最も多かった広告配信枠から、上位50キャンペーン*3の画像広告392種類に関する表3.1のデータを取得した。広告が配信される枠はクリックの有無に大きな影響がある。例えば、あるWebサイトを表示した際に上段の広告配信枠はユーザーの目に止まりやすいが、最下段の広告配信枠はユーザーがスクロールをしなければ表示されず、広告効果としては低いと考えられる。本章では、画像広告の要素に着目していることから広告配信枠を1つに固定した。なお、キャンペーンについても同一条件とすると画像広告の種類が限定されることから、複数キャンペーンを許容することとした。

表 3.1 広告配信ログデータ

変数名		データ型	説明
広告配信日	Day	Categorical	日
広告配信枠	Destination	Categorical	広告が配信された枠の id(1 種類)
キャンペーン	Campaign	Categorical	広告商材や配信対象属性に固有の id

画像広告については Microsoft Computer Vision API(以下、MCV)を利用し、特徴量の抽出を行った。392の画像広告それぞれについて表3.2のデータを取得した。MCVで

*3 広告が訴求するサービスや商品

は、ブランド検出や人物の推定年齢についても実行できるが、分析対象がモバイル広告の画像であることから不相当と判断し採用しなかった。

表 3.2 Microsoft Computer Vision API から取得したデータ

構成要素のグループ	データ型	変数の数	例
category	continuous	23	c_outdoor(0.81),c_people(0.90),...
tag	continuous	210	t_book(0.52),t_woman(0.87),...
clipart	dummy	1	0 or 1
linedrawing	dummy	1	0 or 1
color	dummy	26	fw_red(1),bk_pink(0),...
adult	continuous	1	0.01,...
racy	continuous	1	0.34,...
displaysize	categorical	4	600×500, 300 × 250, ...
format	categorical	3	gif,jpeg,png

MCV の判定結果として category(カテゴリ), tag(タグ) については範囲を (0 – 1) とする連続値が得られる。大きければ当該カテゴリやタグが画像広告の構成要素に含まれる可能性が高い。例えば、ある画像広告について、室外に女性がいると判定されると、c_outdoor と t_woman について高い値となる。MCV では、category(カテゴリ) は図 3.1 のとおり階層構造を伴った 86 分類、tag(タグ) は数千単位で用意されている。本章で使用した 392 の画像広告群では、category が 23 種類、tag が 210 種類が付与された。使用されている色調の前景色 (foreground color), 背景色 (background color) はダミー変数、画像広告の大きさ (displaysize) は 6 水準のカテゴリカルデータ、ファイルフォーマット (format) は 3 水準のカテゴリカルデータが得られる。また、成人向け表現、差別的表現について (0 – 1) の連続値が得られる。ファイルフォーマットが、gif の場合には複数の静止画を順番に表示することで、アニメーション効果を取り入れている画像広告が存在

する。この場合、MCV は複数の静止画に対して判定結果を付与する。例えば、アニメーションの途中で、人物が登場したとしても t_person のタグが付与される。

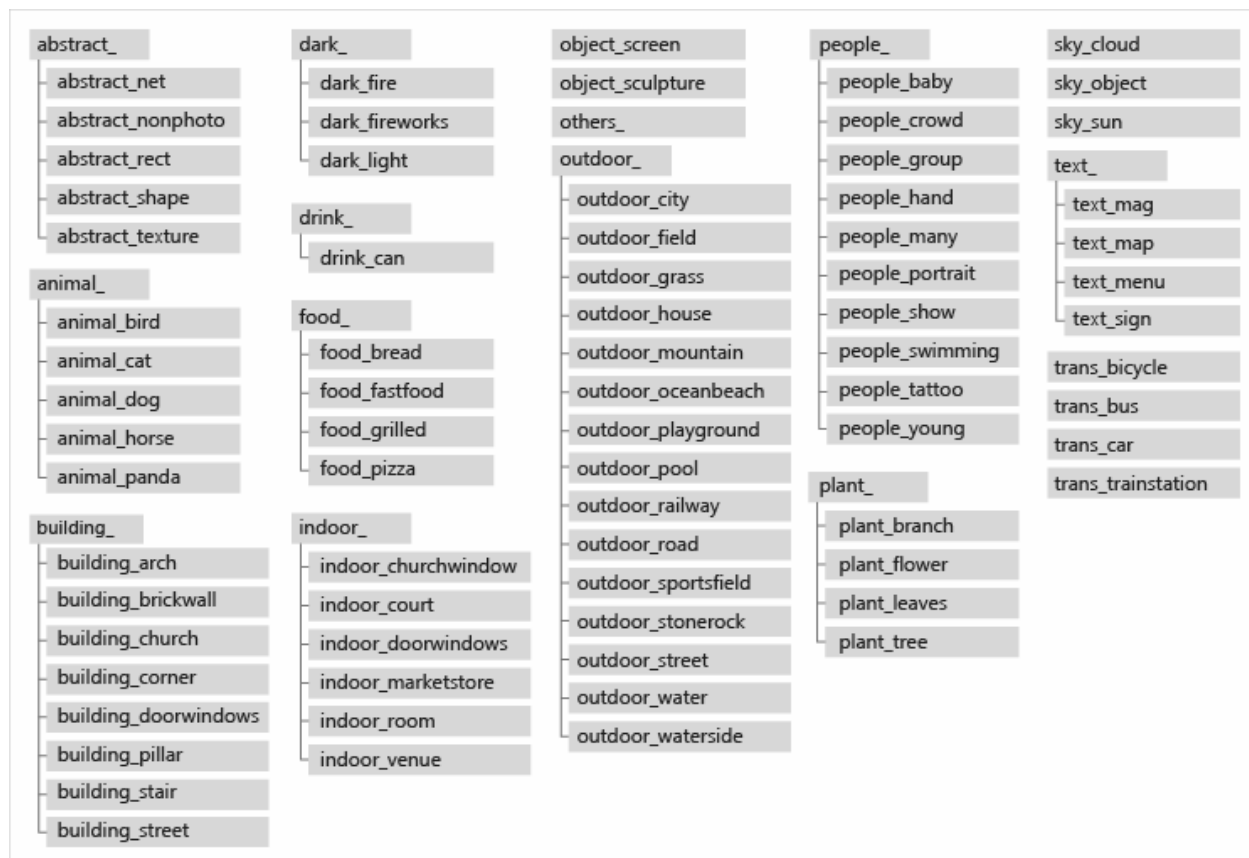


図 3.1 Microsoft Computer Vision API の category 分類 ^{*3}

特徴量のグループは、広告が訴求するサービスや商品 (キャンペーン), 画像広告の構成要素の 2 つに分けられる。表 3.3 で、画像広告数, キャンペーン毎に Impression, Click 数の平均, 中央値, 最大値, 最小値を確認する。画像広告の表示回数の平均は 10,063 回, クリック数は 20 回であった。なお, クリック数が 0 回の画像広告も存在している。キャンペーンにおける表示回数の平均は 78,897 回, クリック数は 153 回であった。

^{*3} [https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-categorizing-images\(2019/7/17\)](https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-categorizing-images(2019/7/17))

表 3.3 基本統計量

データ数			<i>Mean</i>	<i>Median</i>	<i>Max</i>	<i>Min</i>
画像広告数	392	Impression	10,063	1,612	94,123	7
		Click	20	4	185	0
キャンペーン	50	Impression	78,897	18,976	705,023	504
		Click	153	41	1,025	0

3.3.2 と加工

本章では、表 3.4 のように全データを訓練データとテストデータに分割し、訓練データで機械学習によるパラメータ推定を実施後、テストデータで AUC による予測精度の評価を行う。分割後の訓練データ、テストデータともに CTR で全データと大きな差は無い。時間依存性を考慮し訓練データは 6 月 24 日から 6 月 29 日までの 6 日間、テストデータは 6 月 30 日の 1 日とした。

表 3.4 分割したデータ

	画像広告数	キャンペーン数	<i>Impression</i>	<i>Click</i>	<i>CTR</i>
			(表示回数)	(クリック数)	(クリック率)
全データ	392	50	3,944,869	7,650	0.19392%
訓練データ	392	50	2,577,038	5,244	0.20349%
テストデータ	392	49	1,367,831	2,406	0.17590%

CTR 予測問題は、Impression(表示回数)と比較して Click(クリック数)が極端に少ない不均衡なデータを扱う。正例データはクリックされたログ、負例データはクリックされなかったログにあたる。訓練データとして不均衡データをそのまま利用すると、全てクリックしないと予測する学習機が得られる可能性がある。それでも全データを対象した正解率は、表 3.4 から約 99.81%(= 100% - 0.19%)となる。不均衡データを元に機械学習を実行する際には正例データを増やす (Over Sampling), 負例データを減らす (Down

Sampling), 正例を誤答した際のペナルティを重くするなど様々な手法が提案されている [Hayat 19]. 本章では, 計算時間の短縮が期待できる Down Sampling に加えて, [Chawla 02] の Synthetic Minority Oversampling Technique:SMOTE を利用する. 同手法は, 正例を人工的に生成 (Over Sampling) することで過学習を避けつつ, 負例を減らす (Down Sampling) ことができるとされる. [He 14] に習い, 表 3.4 の訓練データについて負例データをランダムサンプリングし, 表 3.5 のように 6 パターンの Down Sampling と全訓練データに対して SMOTE を行った計 7 パターンについて機械学習を行う.

表 3.5 訓練データのパターン

	パターン 1	パターン 2	パターン 3	パターン 4	パターン 5	パターン 6	パターン 7 (SMOTE)
Impression(表示回数)	524,400	104,880	34,960	17,480	10,488	8,740	52,440
Non-Click(非クリック数)	519,156	99,636	29,716	12,236	5,244	3,496	41,952
Click(クリック数)	5,244	5,244	5,244	5,244	5,244	5,244	10,488
CTR(クリック率)	1%	5%	10%	30%	50%	60%	20%

3.4 実証分析

3.4.1 機械学習手法の選択について

本章では, 学習機に Gradient Boosted Decision Trees(GBDT) を利用する. 理由は, 特徴量の重要度の推定が可能であることに加えて, [Deng 19] が提案した Interpretable Trees(inTrees) を GBDT の結果に適用することで, 交互作用についてもルールベースでの解釈ができるからである. GBDT は複数の弱学習機である決定木で構成され, それぞれの決定木から出力された予測が統合されて GBDT モデルとしての予測値となる. Deng(2019) では, 「一つ一つの決定木から全体を理解することは困難であるが, 統合後のモデルを要約することができれば, どのような識別を行なっているか全体を把握できる. また, 各決定木の枝をバスケットとみなし, 統合したモデルの識別ルールをアソシ

エーション分析することで特徴量間における交互作用を抽出できる。」とした。例えば、男性単独の画像広告では効果が低いが、屋外で犬と一緒にであれば効果が高いといったことが考えられることから、単独の効果に加えて交互作用について検討する必要がある。なお、交互作用をモデル化可能な手法としては Factorization Machines[Rendle 10] があり、CTR 予測にも応用されている [Juan 16,Guo 17]。また、[Wang 18] は、Gradient Boosted Decision Trees(GBDT) から交差項の影響をモデル化する手法を提案している。

3.4.2 Gradient Boosted Decision Trees(GBDT)

複数の決定木による分類器をアンサンブル学習において、逐次的に学習し、個々の分類器が苦手とするデータに対して、より分類能力が高くなるように個々の分類器を組み合わせ学習することを勾配ブースティングと呼ぶ。本章では、勾配ブースティングを利用した集団学習手法の一つである Gradient Boosted Decision Trees(GBDT) で予測モデルの学習を行う。以下、[横山 17] に則して GBDT の説明を行う。

勾配ブースティング法は、弱学習器を加法的に追加する学習方法である。非線形な予測モデルである決定木を弱学習器にした勾配ブースティング法は、非線形な予測モデルを構築する。特徴ベクトルを X とした 2 値分類 $y \in \{+1, -1\}$ とする。勾配ブースティング法により最終的に学習される予測モデル f_t は t 本の決定木から構成され、 $f(X) = T_1(X) + T_2(X) + \dots + T_t(X)$ と書ける。ただし、 T_j は 1 本の決定木を示し、実数値を返す。勾配ブースティング法では、以下の式のように損失を最小化する関数 (予測モデル) f^* を学習することを目指す。

$$f^* = \min_f \sum_{i=1}^N \Phi(y_i, f(X_i)) \quad (3.1)$$

ただし、損失関数 Φ は逸脱度で、以下のように書ける。

$$\Phi(y, f(X)) = \ln(1 + \exp(-2yf)) \quad (3.2)$$

3.4.3 AUC による予測精度比較

CTR 予測問題は、広告表示ログそれぞれについてクリックの有無を問う「教師あり」学習である。表 3.5 の通り、Down Sampling と SMOTE による 7 パターンの訓練データで学習を行い、テストデータで AUC による予測精度比較を行う。特徴量は、表 3.1 のキャンペーン、と表 3.2 の MCV による画像広告情報を利用する。学習機には GBDT に加えて、Logistic Regression(LR) を比較対象とした。表 3.6 から AUC は、正例比率を 20% で訓練したパターン 7(SMOTE) の GBDT モデルが最高となった。

表 3.6 各パターンの AUC

Model	パターン 1	パターン 2	パターン 3	パターン 4	パターン 5	パターン 6	パターン 7 (SMOTE)
LR	0.57313	0.57088	0.57404	0.56341	0.56341	0.55304	0.55612
GBDT	0.57544	0.57354	0.57547	0.56219	0.55232	0.54748	0.57951

3.4.4 特徴量の重要度

以下、表 3.6 のパターン 7 の GBDT モデルについて検討する。まず、表 3.7 で特徴量のグループ別に重要度の確認を行う。

表 3.7 画像広告構成要素のグループ別重要度

構成要素のグループ	重要度
キャンペーン	30.62%
category	0.53%
tag	36.78%
color	28.05%
adult	1.32%
racy	2.70%
合計	100%

特定の広告配信枠を対象とした場合に、キャンペーン(広告が訴求するサービスや商品)が大きな影響(30.62%)を持つことが確認できる。画像広告の要素の重要度合計は、約70%となった。内訳は、tag(37%), color(28%), racy(2.7%), adult(1.3%), category(0.5%)となった。更に図3.2, 図3.3で、主な画像広告の構成要素別の重要度について詳細を確認する。

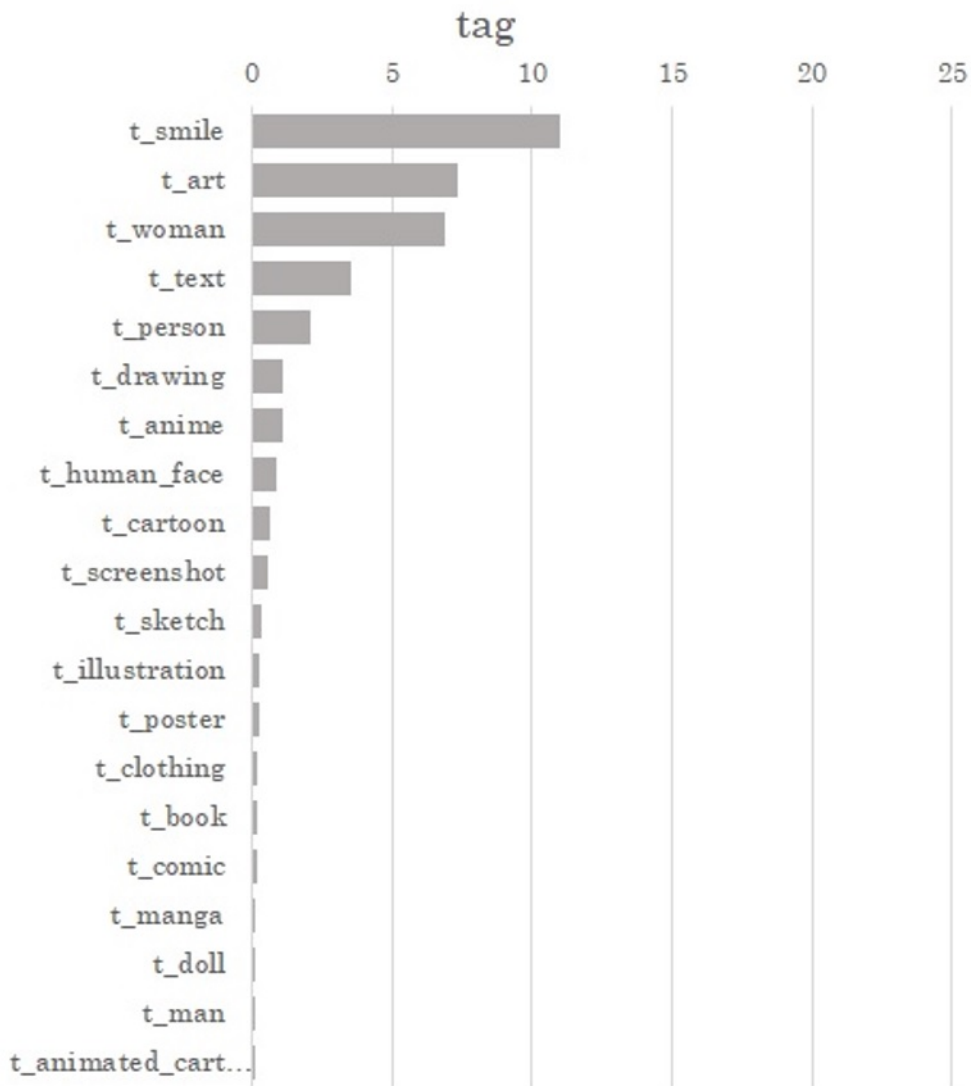


図 3.2 画像広告構成要素の重要度 1

タグ (tag) からは、画像広告内の人物の表情が笑顔 (t.smile:10.9%) である、芸術

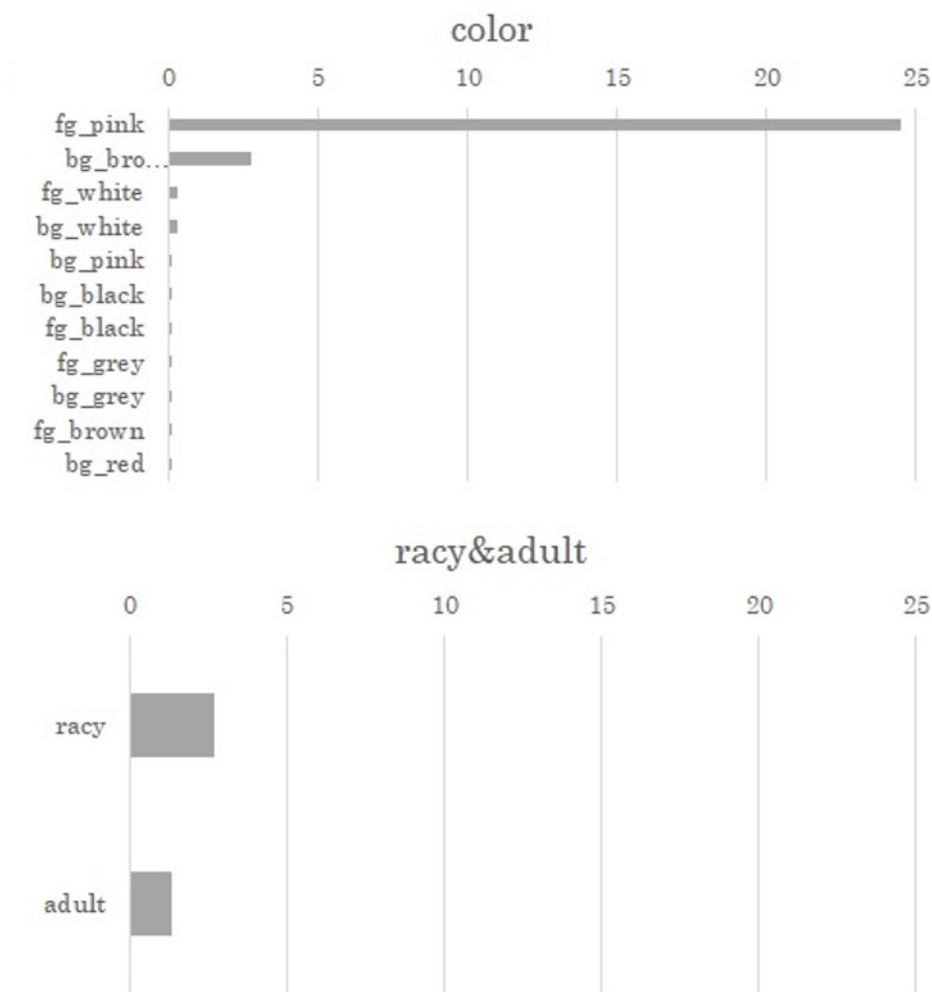


図 3.3 画像広告構成要素の重要度 2

(t_art:7.3%) 的である, 女性 (t_woman:6.8%) が含まれている, テキスト文 (t_text:3.5%) での訴求がある, 人物 (t_person:2.1%) が含まれていることがあげられる. 色彩については, 前景色がピンク (fg_pink:24.5%), 背景色が茶色 (bg_brown:2.8%) が高い結果となった. 差別的/成人向け表現の重要度はそれぞれ racy:2.7%, adult:1.3% となった.

3.4.5 inTrees による交互作用の推定

続いて, Deng(2019) による Interpretable Trees(inTrees) をパターン 7 の GBDT モデルに適用する. inTrees の結果はアソシエーションルールとして得られる. アソシエー

シジョンルールとは、データ内の共起頻度の高い事象の組み合わせをルールとして抽出したものであり、 $\{A \Rightarrow B\}$ の形式で表現される。訓練データから得られた GBDT に対して inTrees を実施、その結果から条件が 2 つのルールについて抜粋^{*4}、テストデータに適用後、下記の値を算出し表 3.8 で示した。Support(支持度) は条件とクリックが同時に観測される確率、Confidence(確信度) はクリックの条件付確率、リフト値は信頼度を期待信頼度で除したものであり条件の有効性を示す。なお、期待信頼度はクリック率 (0.17590%) が相当する。

表 3.8 テストデータにおけるリフト値例

No	クリックの条件	条件該当のレコード数	内クリック数	Support	Confidence	リフト値
1	$t_human_face > 0 \ \& \ t_smile > 0$	58,221	109	0.00797%	0.18722%	1.06435
2	$t_text > 0 \ \& \ t_book > 0$	288,022	510	0.03729%	0.17707%	1.00666
3	$t_smile > 0 \ \& \ t_art > 0$	1,941	3	0.00022%	0.15456%	0.87868

ルールの有効性を確認する。No1 の条件 ($t_human_face \& t_smile$) がテストデータで該当するレコード数は 58,221 件、クリック数は 109 件であり、リフト値は 1.06 と 1 を上回り有効性が確認できた。No2 の条件 ($t_text \& t_book$) もリフト値は 1 を僅かながら上回ったが、No3 の条件 ($t_smile \& t_art$) については下回る結果となった。No3 の条件は訓練データではフィットしたが、テストデータではそれほど適合しなかったものと考えられる。図 3.2, 図 3.3 から画像広告の構成要素の中で、MCV で付与したタグ (tag) 情報や色彩情報 (color) が予測モデルにおいて重要であること、その中でも t_smile , t_art , t_woman , fg_pink , bg_brown などが単独では重要であることが示唆された。また、相互作用を検討するために Deng(2019) で提案された inTrees によって、アソシエーションルールを抽出しテストデータで検証を行った。 t_human_face と t_smile , t_text と t_book , についてはリフト値の上昇が確認できたが、 t_smile と t_art については効果がみられなかった。

^{*4} 条件が 1 つのルールや訓練データに過剰適合 (テストデータに存在しないなど) していると判断されたルールは除外した。

3.5 本章のまとめ

本章では、コンピュータービジョンによる特徴量抽出と、変数の重要度が推定可能な機械学習手法を活用することで、クリックに対して大きな影響を与える要因を、人が解釈可能な文字情報として画像から特定し、その重要度と交互作用を定量化する実証分析を行った。[Chen 16b] や [岩崎 18] では、CTR 予測に画像情報を利用し Saliency Mapping により、予測に寄与した部分を Pixel 単位で可視化し、画像情報と CTR 予測との関連を明らかにしている。Pixel 単位で可視化された情報は有益であるが確認するには時間がかかる可能性が高い。本章では、MCV を利用することで、客観的な文字情報と関連したダミー変数か連続値が特徴量として画像に付与される。よって、意図的に盛り込んだ笑顔の描写 (`t_human_face` と `t_smile`) や色彩 (`fg_pink`) が、MCV で認識できているか画像制作時に確認することができる。また、付与された特徴量は数値データとして扱うことができ、CTR 予測などへの利用も容易である。

インターネット広告の分野での応用として、ユーザー属性の違いに応じて適切な画像広告 (広告クリエイティブ) を出稿できる可能性がある。特定のユーザー属性において効果の高い画像広告要素を意図的に盛り込み、コンピュータービジョンの分析結果から意図したカテゴリやタグ情報が付与されるか客観的な事前評価もできる。また、CTR 予測では画像広告は、付与された固有 id をカテゴリカルデータとして扱われることが多いが、コンピュータービジョンによる特徴量を加えることで、新規の画像広告の予測精度改善が期待される。(一定の表示回数やクリック数に達するまでの期間におけるコールドスタート問題と呼ばれる。) 更なる応用範囲として、複数の装丁パターンがある書籍と「売り上げ」、写真画像と「いいね」数、映画ポスターと「興行収入」など、画像広告にとどま

らず画像から人が解釈可能な特徴量を抽出し，教師データとの関係を探ることができる．

本章の課題として，以下の3点をあげる．まず，予測モデルそのものの改善の余地がある．相互作用を考慮したモデルとして，[Juan 16,Guo 17] で用いられた Factorization Machines や，[Wang 18] で用いられた Tree-enhanced embedding model などがあげられる．次に特徴量に関して，Microsoft Computer Vision に限らず他サービスの利用の検討や，画像広告に埋め込まれた文字情報の活用がある．最後に，訓練データの加工方法があり，特に SMOTE については [Fernandez 18] において近年の研究成果が報告されており検討の余地がある．

第 4 章

アンサンブル学習と LDA の統合による動画広告の要因分析

4.1 はじめに

2019年にインターネット広告費は2兆1,048億円とテレビ広告費の1兆8,612億円を上回った。インターネット広告の中でも、動画広告は前年比157.1%の3,184億円と伸び率が高い。背景として、高機能なスマートフォンが広く普及したこと、動画が画像やテキストと比較してより複雑な情報を短時間で届けられる特徴を持っていることがあげられる。また、2020年に商用利用が開始された5Gサービスでは、「基地局からの下り通信では、最大20Gbpsと、現在の4Gの約100倍の伝送速度となり、例えば2時間の映画を3秒程度でダウンロードできる」とされ[大槻 19]、動画によるコンテンツはより一層身近なものになっていくと考えられる。画像認識分野の研究は、[Krizhevsky 12]がImageNet Large Scale Visual Recognition Challenge (ILSVRC)で、ディープラーニングによる飛躍的な精度向上を果たして以降、関連技術が目覚ましく発展し実務でも広く活用されるようになった。Amazon Web Services, Microsoft AzureやGoogle Cloud Platform(GCP)といったクラウドサービスを提供する大手IT企業は、顧客に対してクラウドに保管されたデータの活用を推進するため、様々な機械学習の分析技術を提供している。これらの

分析サービスを利用することで、ユーザーは自ら大量のテキスト・画像・動画素材を用意して情報を付与するアノテーション作業を行う必要が無く、学習済みモデルを利用することができる。画像・動画素材に関連するタスクであってもクラウドサービスの利用を前提とすると、既にコモディティ化しているといえよう。

インターネット広告における画像広告や動画広告の制作は、クリエイターの勘や経験に頼るところが大きく、データに基づいた効果的な広告クリエイティブの要件（キャラクター、色調、場面展開の数、これらの組み合わせなど）をあらかじめ認識することができれば、大幅に制作工数を削減できる可能性がある。広告主の期待に沿うため短時間でより効果的な広告クリエイティブを制作する必要性から、株式会社電通デジタルや株式会社サイバーエージェントでは、広告クリエイティブの自動生成に関する支援サービスの提供を開始している [株式会社電通デジタル 20, 株式会社サイバーエージェント 20]。

本章の目的は、動画広告効果の要因分析手法を提案することである。動画広告から解釈可能なキーワードを抽出し、広告の効果（コンバージョン）に対して重要な要素を相互作用も考慮した上で特定する。動画広告から構成要素を抽出する際には、コンピュータビジョンを利用する。動画広告に対して GCP サービスの一つである Video Intelligence API（以下、VIA）による解析を実施し、動画広告に含まれる物体検出結果のラベル、シーン数（映像の変わり目を検出し場面転換の数をカウントしたもの）、テキスト情報を抽出し特徴量として利用する。テキスト情報については、Latent Dirichlet Allocation (LDA) [Blei 03, Katsumata 17, 崎濱 19a] によるトピック分布を特徴量とする。得られた結果を、広告効果の有無に関する CVR (Conversion Rate) 予測による 2 値分類問題の特徴量とする。アンサンブル学習とルール抽出を行うことで、特徴量単体の重要度のみならず、相互作用まで考慮することが可能となる。

本章の構成は以下のとおりである。次節では、画像広告と動画広告に関するコンピュー

タービジョンの応用の観点から先行研究のレビューを行う。第3節では、使用した広告配信ログと動画広告データに関して説明する。第4節では、訓練データの加工方法と検証に利用する学習器について、AUC (Area Under Curve) による予測精度比較を行う。次に、得られた学習器とルール抽出の結果から、CVR 予測に対して重要な動画の構成要素とその交互作用について特定する。最後に、第5節において本章の課題を示す。

4.2 先行研究

インターネット広告配信において、広告がクリックされる/されないを扱う CTR (Click Through Rate) 予測問題は収益に直結する重要な課題であることから関連事業を営む Google, Facebook, Yahoo Japan, 株式会社サイバーエージェントなど国内外の大手 IT 企業を中心として膨大な研究が報告されてきた [He 14, McMahan 13, 宮西 14, 田頭 13]。近年はユーザーの行動ログやデバイス情報 (端末機種など) に加えて、画像広告の要素そのものを予測モデルに取り込む研究が進展している。例えば、クリエイターの審美眼を明示的に扱った研究として [石川 19] がある。クリエイターが予め良/悪の教師ラベルを画像広告に対して付与したデータを用いることで、新規画像広告についてもクリエイターの判断を取り入れた良/悪の判定ができる可能性がある。[Zhang 17] では、映画などの動画コンテンツと画像広告の最適なマッチング手法として、深層学習を活用した手法を提案している。[Xia 20] は、画像広告 (バナー広告) を対象に画像、テキスト、デバイスやランディングページ (商品紹介ページ) の URL など複数の情報を組み合わせた特徴量を用いた深層学習による CTR 予測モデルを構築し精度向上を確認している。また、Sensitivity map を用いることで画像の領域の中で広告効果が高い範囲を可視化できるとした。

動画コンテンツの推薦手法として、[Wang 19] は、Hulu Challenge と呼ばれる TV シリーズと映画コンテンツを題材とした推薦システム問題に対して、2つの深層学習モデル

を組み合わせることで比較的良好な CTR 予測が可能になるとした。動画広告の効果予測研究として [Lou 18] は、動画広告を配信する前にその効果を予測するために、深層学習によって抽出した画像に関する特徴量だけでなくテキスト情報を考慮したモデルを提案している。

[Kaminski 20] は、クラウドファンディングサイトを対象に、説明文のテキスト、プレゼンテーション動画の特徴量から、クラウドファンディングの成否を複数の機械学習手法を用いて予測するモデルの提案を行った。動画の音声情報は Google Speech-to-Text API を用いてテキスト情報に変換し、動画は VIA によるタグ付けを行っている。また、特徴量の重要度は罰則付きのロジスティック回帰の結果から得られるが、相互作用については今後の課題とされた。

多くの先行研究において、動画から抽出した特徴量を直接解釈することは意図していないことから、動画広告を制作するクリエイターが理解可能なキーワードやラベルによるフィードバックは難しいことが想定される。画像広告については、[崎濱 19b] が画像広告内の構成要素についてコンピュータービジョンによってラベル付けを行うことで、解釈可能性の高い特徴量を抽出しアンサンブル学習を実施、更に inTrees によるルール抽出の結果から特徴量間の相互作用を考慮できる手法を提案している。

本章では、[崎濱 19b] を動画広告を対象として発展させ、課題として残されていたテキスト情報について考慮する。広告配信ログ、動画広告の構成要素、テキスト情報からなるマルチモーダル的アプローチを行う。テキスト情報の特徴量としてはトピックモデルの結果を用い、機械学習手法とルール抽出方法については、それぞれ [崎濱 19b] の Gradient Boosted Decision Trees(GBDT) と inTrees よりも望ましいと考えられる、XgBoost(eXtreme Gradient Boosting) と defragTrees による手法を提案する。

4.3 使用データと分析対象

4.3.1 使用データ

2019年7月8日から7月14日までの、1週間分のスマートフォン向けのモバイル動画広告配信ログを基に、Impression (表示回数) の多い広告配信枠に配信された63キャンペーンの動画広告230種類に関する表4.1のデータを取得した。動画広告がクリックされた回数がクリック数、動画広告を経由して広告目的が達成された回数がコンバージョン数となる。本章での広告目的はスマートフォン向けアプリのダウンロードである。広告が配信される枠はクリックやコンバージョンの有無に大きく影響する。Webサイト上段の広告配信枠はユーザーの目に留まりやすいが、下段の広告配信枠はユーザーがスクロールをしなければ表示されず、広告効果は低い傾向にある。本章では、動画広告の要素に着目しているため広告配信枠を1つに固定し配信枠の影響を取り除いた。

動画広告についてはVIAを利用して、230の素材それぞれについて、表4.2のデータを取得した。VIAは「保存されている動画とストリーミング動画上の20,000種類を超えるオブジェクト、場所、アクションを自動的に認識します。」とされる[Google 21]。VIAは動画内の著名人の顔を特定する機能や物体の位置や表示されている時間、音声情報の取得など豊富な機能を備えているが、本章で必要と判断した特徴量のみを利用する。[崎濱19b]の課題として残されていたテキスト情報も対象とする。

VIAによる動画広告素材の分析結果として、動画広告内に含まれるシーン数(場面展開が多いとシーン数は多くなり、少ないとシーン数も少なくなる)は連続値、特定された物体に対するラベル、暴力や猥褻などの不適切表現の有無に関する0,1のダミー変数、抽出されたテキスト情報についての文字列が得られる。

表4.3で、動画広告と広告が訴求するサービスや商品(キャンペーン)毎に、クリック数(Click)とコンバージョン数(CV)の平均、中央値、最大値、最小値を確認できる。動

表 4.1 広告配信ログデータ

変数名	Variable Name	変数の数	説明
広告配信日時	time	1	動画広告の配信日時
コンバージョンの有無	CV	1	広告配信成果の有無
キャンペーン	campaign_id	63	広告商材に固有の id
端末の画面サイズ	screen_size	10	320×568, 375 × 812 など
端末の種類	platform_id	3	iPodtouch,iPhone,iPad
端末の向き	orientation	2	縦, 横
動画広告の長さ	play_time	1	秒

表 4.2 Video Intelligence API から取得したデータ

変数名	Variable Name	変数の数	説明
シーン数	shot_num	1	含まれるシーン数
ラベル	label	478	付与されるラベル
不適切表現	explicit	1	不適切表現の有無
テキスト	ocr_text	1	含まれるテキスト情報

表 4.3 基本統計量

	データ数		<i>Mean</i>	<i>Median</i>	<i>Max</i>	<i>Min</i>
動画広告	230	Click	34.9	5.0	1,065	1.0
		CV	0.7	0.0	81	0.0
キャンペーン	50	Click	127.5	41.0	1,065	1.0
		CV	2.7	0.0	81	0.0

画広告のクリック数の平均は 34.9 回，CV 数の平均は 0.7 回であった．CV 数が 0 回の動画広告も確認された．キャンペーンにおけるクリック数の平均は 127.5 回，CV 数の平均は 2.7 回であった．

4.3.2 データの分割と加工

表 4.4 のようにデータセットを訓練データとテストデータに分割する。続いて、訓練データで機械学習によるモデリングを行った後、テストデータで AUC による予測精度評価を行う。本章では広告クリック後にアプリをダウンロードしたか、しなかったかについての予測モデルを構築する。したがってレコード数はクリック数と等しくなる。時間依存性を考慮し、訓練データは 2019 年 7 月 8 日 0:00~7 月 12 日 14:59 まで、テストデータは 2019 年 7 月 12 日 15:00~7 月 14 日 23:59 までとした。全データ、訓練データ、テストデータの何れもコンバージョン率 (CVR:Conversion Rate) でみて大きな差は無い。

コンバージョン (CV) 数 (168 件) は、クリック数 (8,030 件) と比較して極端に少ない (CVR:2.10%)。ここで正例データは CV されたログ、負例データは CV されなかったログである。このようなデータは不均衡データと呼ばれ、訓練データをそのまま利用して CV の有無を教師データとしたモデリングを行うと、どのようなデータに対しても CV しないと予測する学習器が得られる可能性が高い。しかしながら、そのような学習器をテストデータで評価した場合でも、97.9%(= 100% - 2.1%) の正解率となる。このような不均衡データに対する処理方法に、正例データを増やす (Over Sampling)、負例データを減らす (Under Sampling)、正例データを人工的に生成する (SMOTE) など様々な手法が提案されている。以降では、データの加工方法と効果検証は主たる目的ではないことから、計算時間の負荷を考慮し負例データを減らす手法を採用する。

表 4.4 データの分割

	動画広告数	キャンペーン数	Click (クリック数=レコード数)	CV (コンバージョン数)	CVR (コンバージョン率)
全データ	230	63	8,030	168	2.09%
訓練データ	190	58	5,267	110	2.09%
テストデータ	116	42	2,763	58	2.10%

4.4 実証分析

4.4.1 機械学習手法の選択について

特徴量と CV との関係アンサンブル学習器でモデル化し、特徴量の重要度推定を行った後、学習器に対して [Hara 18] で提案された defragTrees を適用しルールベースの解釈を試みる。Boosting は、複数の弱識別器を組み合わせることで、より精度の高い学習器を得るアンサンブル学習手法の一つである。本章では、アンサンブル学習機に GBDT と比較して、スパースデータの処理に強く計算速度が速い XgBoost[Chen 16b] を利用する。

[崎濱 19b] では画像広告を対象に、GBDT と inTrees [Deng 19] を利用し広告効果が高い特徴量やその組み合わせの特定を行っている。[Hara 18] において、inTrees は抽出されたルール間の領域の重複が多いことから解釈が困難になる性質が指摘され、この点を改善した defragTrees が提案されている。同手法は、ランダムフォレストを確率的なモデルとみなして、ベイズ的モデル選択によって単純なルールモデルへと変換することで、より少ないルールで低い予測誤差を期待できるとされる。

抽出されたテキスト情報については、[崎濱 19a] に倣い形態素解析の後、BOW(Bag of Words) 表現に対して LDA を実施し、各動画広告に含まれるテキスト情報のトピック分布を得た。辞書ファイルには新語・固有表現に強い mecab-ipadic-NEologd[佐藤 17,Sato 20] を利用した。

表 4.5 各トピックに出現する単語

topic1	topic2	topic3	topic4	topic5	topic6	topic7	topic8	topic9	topic10
取得	確定	伝説	獲得	海	無料	ゲームプレイ	ポイント	専門	英雄
アプリ	チケット	戦闘	経験	飼育	有料	開発中	ダウンロード	物語	育成
月	冒険	少女	システム	自分	コンテンツ	アニメ	日本野球機構	特技	チーム
生理	無料	敵	戦争	初心者	開始	ネコ	公式ホームページ	大陸	地図
体重	ペット	攻撃	発展	水族	限定	にゃん	公認	料理	戦国

表 4.5 に LDA の結果得られたトピック毎の頻出単語を示す。topic1 は “生理”， “体重” といった単語から体調管理系に関する広告であると考えられ， topic6 では “有料”， “無料”， “コンテンツ”， “限定” といった単語から， 期間限定で有料コンテンツが無料で利用できるキャンペーンについて訴求しているのではないかと考えられる。

4.4.2 XgBoost について

XgBoost の損失関数は以下のように定義される。(1) 式の第一項は観測値と予測値の差を表している。第一項のみでは決定木の特性から過学習が懸念され，汎化性能が低下する可能性が高い。このためペナルティ項として (1) 式の第二項に (2) 式の正則化関数 Ω を組み込んでいる。正則化には $L1$ ノルムや $L2$ ノルムなどが用いられる。 k 番目の決定木を関数 f_k ， x_i を入力値， T を最終ノードの数， w を決定木の葉の重み， γ ， λ をハイパーパラメータとする [Chen 16b,Zhang 18]。

$$\phi(F_t) = \sum_{i=1}^N L(y_i, F(x_i)) + \sum_{k=0}^t \Omega(f_k) \quad (4.1)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (4.2)$$

4.4.3 AUC による予測精度比較

続いて， Down Sampling による 3 パターンの訓練データで学習を行い， テストデータで AUC による予測精度比較を行う。訓練データに含まれる CV 数は 110 回であり， パターン 1 はクリック数 1,100 回， CVR10%。パターン 2 はクリック数 366 回， CVR30%。パターン 3 はクリック数 220 回， CVR50% とした。 Down Sampling においてクリックデータはランダムに選択される。特徴量は， 表 4.1 の広告配信ログデータと表 4.2 の VIA

表 4.6 AUC 比較:訓練データ CVR 別× LDA 有無別

	パターン 1	パターン 2	パターン 3
クリック数	1,100	366	220
CV 数	110	110	110
CVR	10%	30%	50%
AUC (LDA 有)	0.88025	0.87383	0.84043
AUC (LDA 無)	0.87985	0.87068	0.82391

による動画広告の要素を利用し，学習器には XgBoost を用いる．また，テキスト情報として LDA から得られたトピック分布を特徴量に追加する場合 (LDA 有) と，しない場合 (LDA 無) についても比較を行う．

表 4.6 から AUC で最大となったのは，パターン 1(**0.88025**) であった．トピック分布の効果については，除いた場合の AUC，**0.87985** と比較して限定的なものに留まった．

4.4.4 特徴量の重要度比較

以下，表 4.6 のパターン 1 で LDA を利用した XgBoost モデルについて検討する．まず，表 4.7 から特徴量のグループ別に重要度を合計した上で確認を行う．特定の広告配信枠を対象とした場合に，動画広告の抽出したラベルの重要度は，85.93% となった．続いて，テキスト情報が 6.33%，シーン数が 0.96% となり合計 93.22% とコンピュータービジョンによる特徴量がコンバージョン予測に対して支配的な影響を持つことが確認された．端末に関する情報 (画面サイズ，向き，種類) は合計して 4.52% (= 2.75% + 1.55% + 0.22%) と重要度としての影響は限定的であった．

図 4.1，図 4.2 で，主な動画広告の構成要素別の重要度について詳細を確認する．最も重要度の高いラベルは Clip_art (21.250%) であり，挿絵の目的で利用されるイラストのクリップアートを指す．リアリティを追求した動画ではなく，一目で想像上のフィク

表 4.7 動画広告構成要素のグループ別重要度

	構成要素のグループ	重要度
VIA _(Video Intelligence API)	ラベル	85.93%
	テキスト情報	6.33%
	シーン数	0.96%
端末情報	端末の画面サイズ	2.75%
	端末の向き	1.55%
	端末の種類	0.22%
配信ログ	キャンペーン	1.79%
	動画広告の長さ	0.48%
	合計	100%

ションであると認識できるような内容であることが窺われる。2,4位は、それぞれ braid (18.694%) , long_hair (6.050%) と髪型に関するラベルが並んでいる。braid は三つ編みを指すことから三つ編みや長髪の女性が登場した動画広告の広告効果が高かったことが示唆されている。3位は text (6.109%) であり、動画広告内のテキスト情報の重要度が比較的高いことが確認された。その他は、song (5.442%) , gameplay (4.748%) , fiction (3.957%) , emotion (3.851%) と続き、歌や実際のゲーム画面、創作物としての映像表現、感情に訴えかける動画要素の重要度が並んだ。

図 4.3 に各トピックの重要度を示した。最も高いトピックは topic3(1.543%) , 続いて topic7(1.534%) であった。表 4.5 から topic3 の出現単語【“伝説”, “戦闘”, “少女”, “敵”, “攻撃”】からロールプレイングゲームが想起される。topic7 の出現単語【“ゲームプレイ”, “開発中”, “アニメ”, “ネコ”, “にゃん”】からは、これからリリースされるアニメのネコが活躍するゲームであることが窺われる。

4.4.5 defragTrees による XgBoost モデルの解釈

次に、XgBoost の学習器に対して [Hara 18] による defragTrees を適用する。defragTrees の結果はアソシエーションルールとして得られる。アソシエーションルールとは、データ内の共起頻度の高い事象の組み合わせをルールとして抽出したものであり、 $A \Rightarrow B$ の形式で表現される。訓練データから得られた XgBoost モデルに対して defragTrees を実施、その結果からコンバージョンがあった場合のルールについて抜粋、テストデータに適用し表 4.8 を得た。Support (支持度) は条件とクリックが同時に観測される確率、Confidence (確信度) はクリックの条件付確率、リフト値は信頼度を期待信頼度で除したものであり条件の有効性を示す。なお、期待信頼度はテストデータの CVR (2.10%) が相当する。

No.1 の条件 ($animal \geq 1 \& text \geq 1 \& shot_num < 2$) がテストデータで該当するレコード数は 243 件、CV 数は 13 件、リフト値は 2.551 と 1 を大きく上回り有効性が確認できた。コンピュータービジョンで動物と判定される描写がテキストメッセージとともに表現され、動画内の場面転換が「無い」か「1 回」の動画広告の効果が高いことが示唆された。No.2 の条件 ($drawing \geq 1$) もリフト値は 3.905 と 1 を大きく上回った。drawing については変数重要度ではラベルの上位 20 位以内に入っていないが defragTrees では重要なルールであると判断され、リフト値でも有用性が確認された。線描で描かれた動画広告効果の高いことが窺われる。No.3 の条件 ($LDA_topic_9 \geq 0.5 \& LDA_topic_10 = 0 \& shot_num < 2$) もリフト値 (2.109) から有効性が示唆された。表 4.5 から topic9 の【“特技”, “専門”, “大陸”, “料理”, “物語”】といった単語が含まれ、topic10 の【“英雄”, “育成”, “チーム”, “地図”, “戦国”】が含まれない、かつ、場面転換が「無い」か「1 回」の動画広告の効果があることが窺われた。図 4.3 においてトピック分布の重要度を示しているが、topic9 は 0.692%, topic10 は 0.383% と単独での重要度は他の特徴量と比較して高くはないが、交

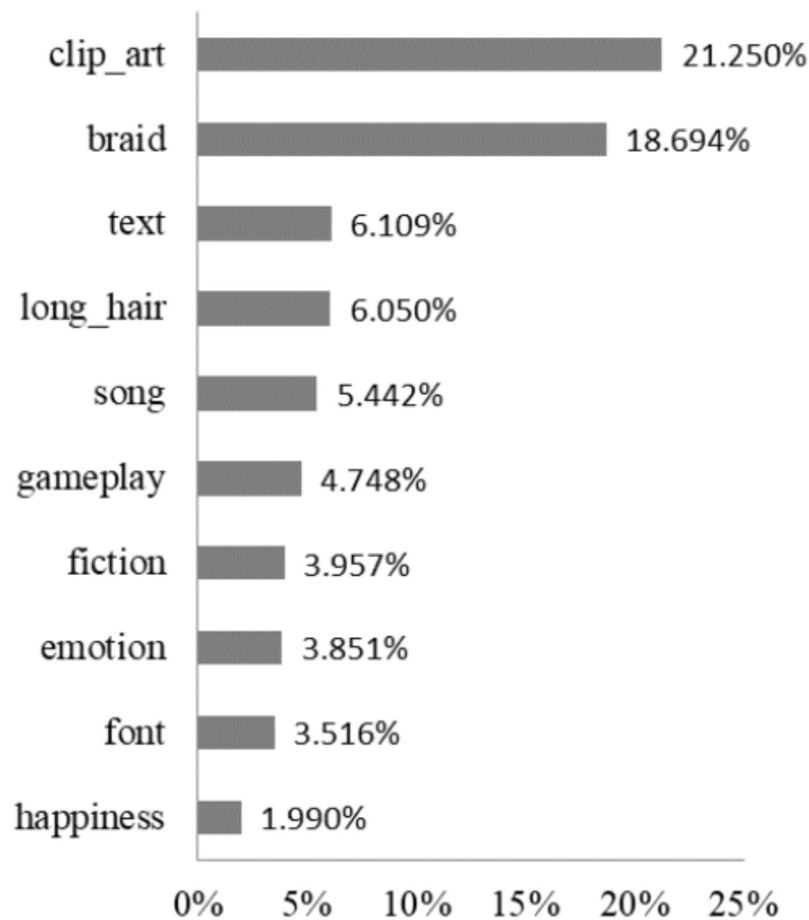


図 4.1 動画広告構成要素の重要度 1

相互作用を考慮すると注目すべき特徴量であることが確認できた。

4.5 本章のまとめ

本章では、動画広告効果の要因分析を目的とし、コンピュータービジョンを活用して人が解釈可能なラベルを動画内の物体に自動的に割り当て、更にテキスト情報を抽出した後、広告配信のログ情報に加えて、割り当てられたラベルとテキスト情報 (LDA によるトピック分布) を特徴量とし、コンバージョン (CV) の有無を教師データとした XgBoost によるモデル構築を行い変数重要度を推定した。最後に、得られたモデルに対して defragTrees による相互作用を考慮したルール抽出を行った。

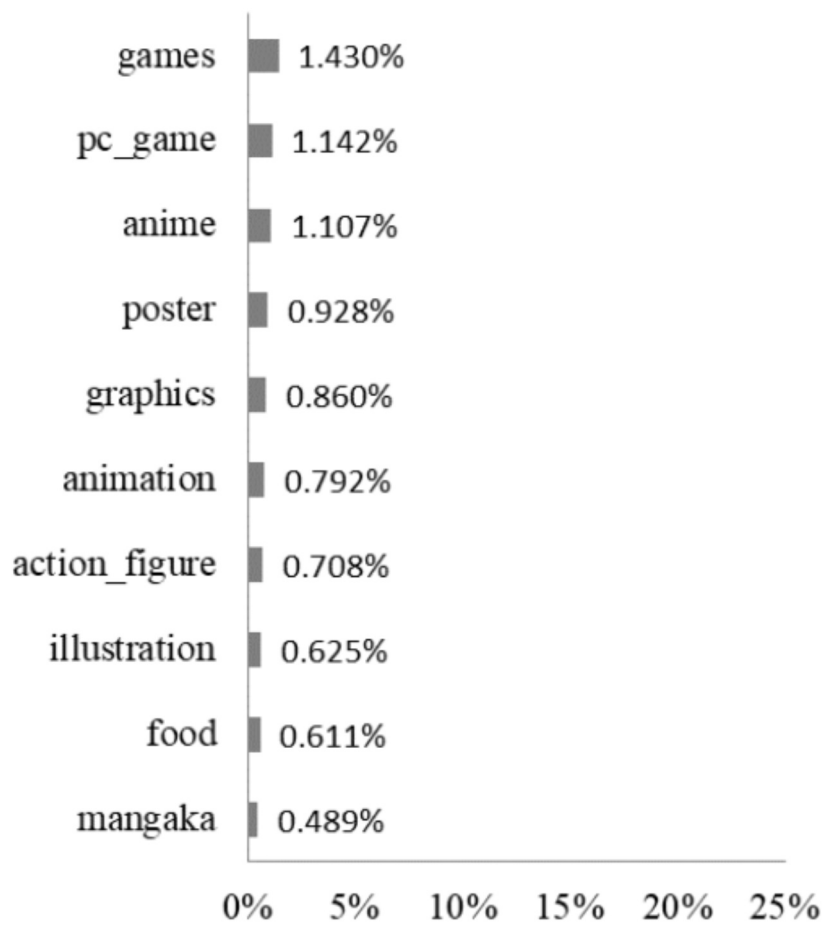


図 4.2 動画広告構成要素の重要度 2

表 4.8 テストデータにおけるリフト値例

No	CV 条件	条件該当の レコード数	内 CV 数	Support	Confidence	リフト値
1	$animal \geq 1 \ \& \ text \geq 1$ $\& \ shot_num < 2$	243	13	0.471%	5.350%	2.551
2	$drawing \geq 1$	403	33	1.194%	8.189%	3.905
3	$LDA_topic_9 \geq 0.5 \ \& \ LDA_topic_10 = 0$ $\& \ shot_num < 2$	294	13	0.471%	4.422%	2.109

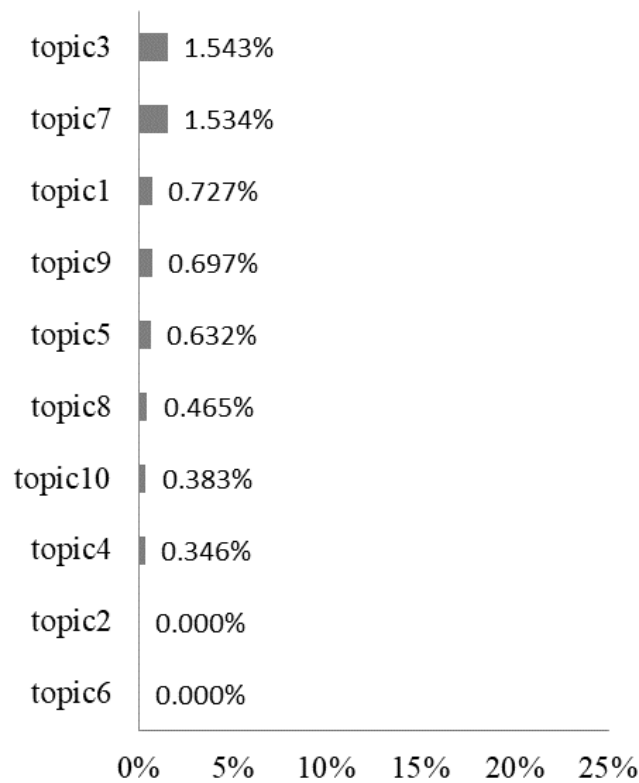


図 4.3 動画広告構成要素の重要度 3(トピック)

コンピュータービジョンを用いて様々な映像コンテンツから特徴量を抽出することが可能であり、その効果についての教師データが取得できれば交互作用を含めたうえで関係を捉えられる。提案手法は汎用性が広く、用途として、映像コンテンツの途中離脱要因の特定、デジタルサイネージの効果向上、映画やテレビ番組の予告映像の効果検証などがあげられる。

本章の課題として、大きく3点あげられる。第1に、本章で用いたVIAは特定の外資系企業が提供する学習済みのモデルであるがゆえに、ラベル付けを実行する際に日本特有の表現や広告ならではの特徴を把握し切れていない可能性がある。画像広告、動画広告ともに専用の学習用素材を用意し、アノテーション作業を実施することで、より精度の

高い学習済みモデルが得られると期待される。アノテーション作業が困難な場合の代替案としては、国内外複数のクラウドサービスを用いた物体検出によるラベル付けを実行することで同問題点を緩和できる可能性がある。第2に、音声情報の活用がある。音声には人が解釈可能なテキスト情報の他に、感情に訴えかける音楽、各種効果音が盛り込まれている。映像、テキスト、音声を含めたマルチモーダル学習の検討の余地がある。最後に、有用性の観点からクリエイターへのフィードバックの質を高める研究があげられる。分析の結果得られたルールを踏まえた広告素材による効果検証を行い、クリエイターの立場から特徴量やルール抽出の改善点に関する意見を集め、人とテクノロジーの協業を進化できるものと期待できよう。

第5章

おわりに

5.1 本研究のインプリケーション

第3章では画像広告を，第4章では動画広告を対象としてアンサンブル学習と機械学習の解釈手法から相互作用を含めた効果検証の方法を提案した．広告配信のログに関しては情報に差はなく，それぞれの広告素材に応じた特徴量の検討を行った．なお，インターネット広告において広告が配信される広告掲載場所は広告効果と密接に関連している．例えば，Webサイトの最上段であれば表示されて人の目に触れる可能性が高い一方，Webサイトの最下段ではスクロールを続けなければそもそも表示されないため，広告の素材がどれだけ優れていてもその効果は限定的となる．このため，本研究においては特定の掲載場所に表示された広告を対象とすることで広告素材の違いによる広告効果を測定した．画像広告と動画広告に関して，コンピュータービジョンによる特徴量抽出の結果を比較すると，画像広告で利用できる特徴量の数は動画広告より格段に少ない．そのため，第3章では色彩に関する特徴量を利用し，その構成要素のグループ重要度は，28.05%とtag，キャンペーンに続いて3番目となった．一方，動画広告については含まれる人，動物，物のオブジェクトに加えて，オブジェクトの位置情報も含めると取得可能な特徴量が膨大である．このため，人が解釈してクリエイティブ制作に活用しやすい特徴量を選択した．具体的には，動画に含まれるオブジェクトのラベルの他に，場面転換を表すシーン

数、動画広告の長さ、広告内で表示されたテキスト情報を利用した。

予測精度の改善が目的であれば利用できる特徴量は多ければ多いほど望ましく、[Chen 16a]のように、CTR 予測モデル構築の際に、広告配信ログに加えて画像広告のピクセル情報も Deep Learning の特徴量として利用する方法がある。また、説明可能 AI の技術を利用して、画像広告のどの部分が CTR 予測に対して貢献しているか視覚情報として提供することができる。ただし、同手法によるフィードバックをクリエイターが理解する際には、画像広告を一枚ずつ目視で確認して広告効果への貢献が大きい領域に何が描かれているか判断する必要がある。

本研究では、広告内に何が表現されているか、人が判断する代わりにコンピュータービジョンによってキーワードとして取得することを提案した。研究の背景として、現在外資 IT 大手を筆頭に画像や動画を対象としたコンピュータービジョンサービスが展開されており、ビジネスにおいて如何に活用していくかというフェーズに突入しているため、本研究の示唆についても実現性が高いものであるといえる。更に、コンピュータービジョンによる物体検出の精度向上が必要となれば、人手によるサポートを活用することも可能であり、例えば Amazon の Mechanical Turk、国内企業の ABEJA の ABEJA Platform などがあげられる。どちらも、Annotation 作業(人手による教師データ作成)を行うプラットフォームであり、画像広告や動画広告に特化した物体検出モデルの構築も構想できる。また、第 4 章では動画広告内のテキスト情報を抽出して LDA を実行し、得られたトピック分布の結果についても特徴量として利用した。トピックモデルについては、各トピックにおいて出現確率の高いキーワードを得ることができることから広告効果の高いと期待されるキーワードについて特定することが可能である。

インターネット広告はマーケティング活動の手段の一つであり、広告をクリックするとアプリストアや商品詳細に関するランディングページ (Landing Page:LP) などに遷移し、

アプリのダウンロードや商品購入、サービス申込みといったコンバージョンへ誘導することが目的とされる。本研究の応用として、これら広告クリック後の遷移先の Web サイトが画像広告、動画広告のクリエイティブとマッチしているかどうかについても示唆が得られる可能性がある。単純なコンバージョン数や Web サイトの滞在時間だけではなく、Web サイトのクリエイティブについてコンピュータービジョンによる物体検出や自然言語処理によるテキスト分析を行い特徴量として利用することも考えられよう。

5.2 本研究の貢献と限界

本研究では、「人が解釈可能なキーワード」をコンピュータービジョンによって画像広告・動画広告から抽出する、説明可能 AI の中で大域的な説明とされる弱学習器に決定木を用いたアンサンブル学習モデルを相互作用を含めたルールベースで解釈する。効果的な広告クリエイティブを効率的に制作できる手法であり、現時点で実装可能な「人と機械の協業」の事例を示した。

実装した場合の具体的な貢献として、クリエイターが高い広告効果が見込まれる広告クリエイティブを効率的に制作できることから、広告効果の向上による売り上げ増加と工数削減によるコストカットが期待できる。広告配信前に期待される予測 CTR・予測 CVR を推定することも可能である。また、意図的に盛り込んだ物体がコンピュータービジョンによって認識できるか客観的に判定される点も実務上有益であると考ええる。

取り上げた手法は研究が日々進んでおり、改善された機械学習、自然言語処理、コンピュータービジョン、説明可能 AI を用いることでより一層精度の高い結果が得られる可能性が高い。この点は、本研究の限界であるが、これら手法を組み合わせるスキームは「人と機械の協業」において様々な分野に応用可能であると考ええる。

5.3 今後の展望

[電通 20] の日本の広告費によると，国内におけるインターネット広告市場は1兆1,594億円(2015年)から2兆2,290億円(2020年)と過去5年間で2倍の成長を遂げた．2021年7月～9月期決算において広告収入が主な事業である Alphabet(Google) の売上は約7兆4,000億円，利益は約2兆1,500億円，同期の Meta(Facebook) の売上は約3兆3,000億円，利益は約1兆円とコロナ禍においても好調な業績であった．一方で，プライバシー規制を背景にインターネット広告は大きな転換期を迎えている．

代表的なプライバシーに関する規則としては，EU圏におけるデータ保護規制の「General Data Protection Regulation:GDPR」，カリフォルニア州消費者プライバシー法の「California Consumer Privacy Act:CCPA」，日本の2020年に改正された「個人情報保護法」(2022年4月から全面施行)がある．プライバシー規制と関連する重要な技術的な観点に，3rd party cookieの取り扱いがあげられる．3rd party cookieとは，主にインターネット広告の効果測定やリターゲティング広告(Webサイトを1度訪れたユーザーを離脱後も追跡する広告)などで活用される技術であるが，プライバシー規制の強化を背景として Apple の iOS や Google の Chrome から排除される動きがある．3rd party cookie に代わる様々な技術開発も行われているが，「第2章 インターネット広告とその研究の変遷」でも示した通り，本質的にユーザーに必要とされない不快な体験は回避されることになるであろう．なお，現時点においてもアドブロックや広告ブロック(Ad Blocking)と称されるソフトウェアを利用することで，インターネット広告をユーザーが主体的に排除できる．

[戸田 10b] では，「今後の広告活動においては，その広告効果は広告を掲載するメディアの集客力によって決まるものではなく，広告メッセージそのものがコンテンツとしてどれだけの魅力を持っているかによって決定されるものとなっていくであろう」と

結論付けている。

例えば、一部の熱心なファンに愛される広告コンテンツが SNS を通じて拡散することで、興味を持たない一般ユーザーに対しても認知を広げた事例に K-POP がある。K-POP においては、新曲リリース前に「Teaser」(ティーザー) と呼ばれる短い動画広告を戦略的に数本 SNS 上で発表し、新曲に関する情報を少しずつ明らかにすることでファン心理を煽り、期待感を高めるプロモーションが広く実施されている [田中 21]。ファンによる極端な事例として片付けるのではなく、魅力的なコンテンツであればインターネット広告においてもユーザーによってコンテンツが拡散されることで、商品やサービスの魅力を伝えられる一例としても捉えられよう。

データはあくまで過去の事実であり、データが存在しない場合でも人は想像力を元に仮説を立てることができる。テクノロジーの力を援用しつつ、想像・創造・発想など機械よりも人が得意な領域に注力すべきであろうと考える。人間と機械の協調について研究し「サイバネティックス」を提唱した Wiener は、「コンピューターの性能が高くなるほど、創造力と独創力に富む高度な人間の思考とコンピューターを結びつけることが重要になる」としたとされる [Isaacson 19]。本研究が、「人と機械が協業」することによってより魅力的な広告コンテンツを効率的に開発することに貢献することができれば幸いである。

参考文献

- [Amit 94] Amit, Y. and Geman, D.: Randomized inquiries about shape: An application to handwritten digit recognition, *Defense Technical Information Center document* (1994)
- [Angelino 17] Angelino, E., Larus-Stone, N., Alabi, D., Seltzer, M. and Rudin, C.: Learning Certifiably Optimal Rule Lists for Categorical Data, *arXiv preprint*, 1704.01701 (2017)
- [Animesh 11] Animesh, A., Viswanathan, S. and Agarwal, R.: Competing “Creatively” in Sponsored Search Markets: The Effect of Rank, Differentiation Strategy, and Competition on Performance, *Information Systems Research*, Vol. 22, No. 1, pp. 153–169 (2011)
- [Bach 15] Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K., R. and Samek, W.: On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation, *PloS One*, Vol. 10, No. 7 (2015)
- [Bien 11] Bien, J. and Tibshirani, R.: Prototype selection for interpretable classification, *Annals of Applied Statistics website*, Vol. 5, No. 4, pp. 2403–2424 (2021)
- [Blei 03] Blei, D. M., Ng, A. Y. and Jordan, M. I.: Latent dirichlet allocation, *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022 (2003)
- [Borisova 17] Borisova, O. and Martynova, A.: Comparing the Effectiveness of Outdoor Advertising with Internet Advertising, *JAMK University of Applied Sciences* (2017)
- [Breiman 01] Breiman, L.: Random forests, *Machine Learning*, Vol. 45, No. 1, pp. 5–32 (2001)
- [Breiman 96] Breiman, L. and Shang, N.: Born again trees, *Berkeley Technical Report* (1996)
- [Brynjolfsson 13] Brynjolfsson, E. and McAfee, A., 村井章子訳: 機械との競争 (原題: Race Against The Machine), 日経 BP (2013)

- [Çakmak 19] Çakmak, T., Tekin, A., Şenel, C., Çoban, T., Uran, Z. and Sakar, C.: Accurate prediction of advertisement clicks based on impression and click-through rate using extreme gradient boosting, *In Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*, pp.621-629(2019)
- [Chawla 02] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P.: SMOTE: Synthetic minority over-sampling technique, *Journal of Artificial Intelligent Research*, Vol.16, pp.321–357(2002)
- [Chen 16a] Chen, J., Sun, B., Li, H., Lu, H. and Hua, X.: Deep CTR Prediction in Display Advertising, *Proceedings of the 24th ACM international conference on Multimedia*, pp.811-820(2016)
- [Chen 16b] Chen, T. and Guestrin, C.: XgBoost: A scalable tree boosting system, *Proceedings of 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785–794(2016)
- [Cho 04] Cho, H.C. and Cheon, J.H.: Why Do People Avoid Advertising on the Internet?, *Journal of Advertising*, Vol.33, No.4, pp 89-96(2004)
- [Crain 19] Crain, M. and Nadler, A.: Political Manipulation and Internet Advertising Infrastructure, *Journal of Information Policy*, Vol.9, pp.370-410(2019)
- [D'andréa 19] D'andréa, C. and Mintz, A.: Studying the Live Cross-Platform Circulation of Images With Computer Vision API: An Experiment Based on a Sports Media Event, *International Journal of Communication*, Vol.13, pp.1825-1845(2019)
- [Deng 19] Deng, H.: Interpreting tree ensembles with inTrees, *International Journal of Data Science and Analytics*, Vol.7, pp.277-287(2019)
- [Dietterich 02] Dietterich, T., G.: Ensemble learning. In *The handbook of brain theory and neural networks*, Vol.2, pp.110–125, *MIT Press*(2002)

- [Fernández 18] Fernández,A.,Garcia,S.,Herrera,F.and Chawla,N.V.:Smote for learning from imbalanced data:Progress and challenges marking the 15-year anniversary,*Journal of Artificial Intelligence Research*,Vol.61,pp.863-905(2018)
- [Freund 95] Freund,Y.and Schapire,R.,E.:A decision-theoretic generalization of on-line learning and an application to boosting,*In European conference on computational learning theory*,pp.23-37(1995)
- [Frey 13] Frey,C.,B.and Osborne,M.:The Future of Employment:How susceptible are jobs to computerisation?,*Technological Forecasting and Social Change*,Vol.114,pp.254-280(2013)
- [Friedman 01] Friedman,J.H.:Greedy function approximation:A gradient boosting machine,*Annals of Statistics*,Vol.29,pp.1189–1232(2001)
- [Friedman 02] Friedman,J.H.:Stochastic gradient boosting,*Computational Statistics & Data Analysis*,Vol.38,No.4,pp.367–378(2002)
- [Ghose 09] Ghose,A.and Yang,S.:An Empirical Analysis of Search Engine Advertising:Sponsored Search in Electronic Markets,*Management Science*,Vol.55,No.10,pp.1605-1622(2009)
- [Google 21] Video AI,<https://cloud.google.com/video-intelligence?hl=ja>(accessed 2021/2/5)
- [Guo 17] Guo,H.,Tang,R.,Ye,Y.,Li,Z.and He,X.:DeepFM:a factorization-machine based neural network for CTR prediction,*arXiv preprint*,pp.1703.04247(2017)
- [Hara 18] Hara,S.and Hayashi,K.:Making tree ensembles interpretable:A Bayesian model selection approach,*Proceedings of the 21st International Conference on Artificial Intelligence and Statistics (AISTATS'18)*,pp.77-85(2018)

- [He 14] He,X.,Pan,J.,Jin,O.,Xu,T.,Liu,B.,Xu,T.,Shi,Y.,Atallah,A.,Herbrich,R.,Bowers,S.and Quinonero,C.,J.:Practical lessons from predicting clicks on ads at facebook,*In Proceedings of the Eighth International Workshop on Data Mining for Online Advertising,ADKDD'14*,pp.1–9(2014)
- [Ho 95] Ho,T.,K.:Random decision forests,*In 1995 Proceedings of the third international conference on document analysis and recognition:IEEE*,Vol.1,pp.278–282(1995)
- [Isaacson 19] Isaacson,W. 井口耕二訳, イノベーターズ 1, 講談社 (2019)
- [Jansen 08] Jansen,B.,J.and Mullen,T.:Sponsored search: an overview of the concept,history,and technology,*International Journal of Electronic Business*,Vol.6,No.2,pp.114–131(2008)
- [Jiang 11] Jiang,T.,Li,J.,Zheng,Y.and Sun,C.:Improved bagging algorithm for pattern recognition in UHF signals of partial discharges,*Energies*,Vol.4,No.7,pp.1087–1101(2011)
- [Juan 16] Juan,Y.,Zhuang,Y.,Chin,W.S.and Lin,C.J.:Field-aware factorization machines for CTR prediction,*In Proceedings of the 10th ACM Conference on Recommender Systems*,pp.43–50(2016)
- [Kaminski 20] Kaminski,J.C.and Hopp,C.:Predicting outcomes in crowdfunding campaigns with textual,visual and linguistic signals,*Small Business Economics*,Vol.55,pp.627–649(2020)
- [Katsumata 17] Katsumata,S.,Motohashi,E.,Nishimoto,A.and Toyosawa,E.:The contents-based website classification for the Internet advertising planning:An empirical application of the natural language analysis,*The Review of Socionetwork Strategies*,Vol.11,No.2,pp.129–142(2017)
- [Kim 16] Kim,B.,Khanna,R.and Koyejo,O.,O.:Examples are not enough learn to

criticize! criticism for interpretability,*Proceeding in Neural Information Processing Systems(NIPS)*,pp.2280-2288(2016)

[Koh 20] Koh,P.,Wei.and Liang,P.:Understanding Black-box Predictions via Influence Functions,*Proceedings of the 34th International Conference on Machine Learning(PMLR)*,Vol.70,pp.1885-1894(2017)

[Krizhevsky 12] Krizhevsky,A.,Sutskever,I.and Hinton,G.:Imagenet classification with deep convolutional neural networks,*Advances in Neural Information Processing systems*,pp.1097-1105(2012)

[Lage 18] Lage,I.,Ross,A.,S.,Kim,B.,Gershman,S.,J.and Doshi-Velez,F.:Human-in-the-Loop Interpretability Prior,*In Advances in Neural Information Processing Systems*(2018)

[Lakkaraju 16] Lakkaraju,H,Bach,S.,H.and Leskovec,J.:Interpretable decision sets:A joint framework for description and prediction,*In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. (ACM)*,pp.1675-1684(2016)

[Lou 18] Lou,Q.,Sarkhel,S.,Mitra,S.and Swaminathan,V.:Content-based effectiveness prediction of video advertisements,*IEEE International Symposium on Multimedia (ISM)*,pp.69-72(2018)

[Lundberg 17] Lundberg,S.,M.and Lee,S.,I.:A unified approach to interpreting model predictions,*Proceeding in Neural Information Processing Systems(NIPS)*,pp.4768–4777(2017)

[Mcmahan 13] McMahan,H.,B.,Holt,G.,Sculley,D.,Young,M.,Ebner,D.,Grady,J.,Nie,L.,Phillips,T.,Davydov,E.,Golovin,D.,Chikkerur,S.,Liu,D.,Wattenberg,M.,Hrafinkelsson,A.,M.,Boulos,T.and Kubica,J.:Ad click prediction:a view from the trenches,*In Proceedings of the 19th*

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD),pp.1222–1230(2013)

[Meinshausen 10] Meinshausen,N.:Node harvest,*The Annals of Applied Statistics*,Vol.4,No.4,pp.2049–2072(2010)

[Moneera 21] Moneera,A.,Maram,A.,Azizah,A.,AlOnizan,T.,Alboqaytah,D.,Aslam,N.and Khan,U.,I.:Click through rate effectiveness prediction on mobile ads using extreme gradient boosting,*Computers,Materials Continua*,Vol.66,No.2,pp.1681–1696(2021)

[Munawar 19] Munawar,H.,M.,Khan,S.,Zamir,W.,Shen,J.and Shao,L.:Max-margin Class Imbalanced Learning with Gaussian Affinity,*arXiv preprint*,pp.1901.07711(2019)

[Narita 19] Narita Yusuke,Yasui Shota and Yata Kohei:Efficient Counterfactual Learning from Bandit Feedback,*In The 35th AAAI Conference on Artificial Intelligence (AAAI)*(2019)

[Nihel 13] Nihel,Z:The Effectiveness of Internet Advertising through Memorization and Click on a Banner,*International Journal of Marketing Studies*,Vol.5,No.2,pp.93-101(2013)

[Polikar 06] Polikar,R.:Ensemble based systems in decision making,*IEEE Circuits and Systems Magazine*,Vol.6,No.3,pp.21–45(2006)

[Previte 98] Previte,J.and Forrest,E.:Internet Advertising:An Assessment of Consumer Attitudes to Advertising On the Internet,*Australia-New Zealand Marketing Academy Conference*,pp.2040-2052(1998)

[Rätsch 01] Rätsch,G.,Onoda,T.and Müller,K.,R.:Soft margins for AdaBoost,*Machine Learning*,Vol.42,No.3,pp.287–320(2001)

[Reis 18] Reis,A.,Paulino,D.,Filipe,V.and Barroso,J.:Using Online Artificial Vision Ser-

vices to Assist the Blind - an Assessment of Microsoft Cognitive Services and Google Cloud Vision, *WorldCIST*, Vol.746, No.12, pp.174–184(2018)

[Rendle 10] Rendle, S.: Factorization machines, *In 2010 IEEE International Conference on Data Mining*, pp.995-1000(2010)

[Ribeiro 16] Ribeiro, M., T., Singh, S. and Guestrin, C.: "Why Should You trust?" Explaining the Predictions of Any Classifier, *SIGKDD*(2016)

[Ribeiro 18] Ribeiro, M., T., Singh, S. and Guestrin, C.: Anchors: High-precision model-agnostic explanations, *In Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI' 18)*(2018)

[Richards 18] Richards, D.R. and Tunçer, B.: Using image recognition to automate assessment of cultural ecosystem services from social media photographs, *Ecosystem Services*, Vol.31, pp.318–325(2018)

[Sagi 18] Sagi, O. and Rokach, L.: Ensemble learning: A survey, *Data Mining and Knowledge Discovery*, Vol.8, No.4, e1249(2018)

[Sato 20] Sato, T.: mecab-ipadic-neologd(Online) <https://github.com/neologd/mecab-ipadic-neologd>(accessed 2021/2/5)

[Schlosser 00] Schlosser, E., A., Shavitt, S. and Kanfer, A.: Survey of Internet users' attitudes toward Internet advertising, *Journal of Interactive Marketing*, Vol.13, No.3, pp.34-54(2000)

[Schroff 15] Schroff, F., Kalenichenko, D. and Philbin, J.: FaceNet: A Unified Embedding for Face Recognition and Clustering, *arXiv preprint*, 1503.03832v3(2015)

[Shrikumar 18] Shrikumar, A., Greenside, P. and Kundaje, A.: Learning important features through propagating activation differences, *arXiv preprint*, 1704.02685(2017).

- [Sivan 16] Sivan,S.and Darsan,G.:Computer vision based assistive technology for blind and visually impaired people,*In Proceedings of the International Conference on Computing Communication and Networking Technologies*(2016)
- [Smilkov 17] Smilkov,D.,Thorat,N.,Kim,B.,Viégas,F.and Wattenberg,M.:SmoothGrad:removing noise by adding noise,*arXiv preprint:1706.03825*(2017)
- [Springenberg 14] Springenberg,J.,T.,Dosovitskiy,A.,Brox,T.and Riedmiller,M.:Striving for simplicity:The all convolutional net,*arXiv preprint,1412.6806*(2014)
- [Sundararajan 17] Sundararajan,M.,Taly,A.and Yan,Q.:Axiomatic attribution for deep networks,*arXiv preprint,1703.01365*(2017)
- [Surowiecki 04] Surowiecki,J.:The Wisdom of Crowds:Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business,Economies,Societies and Nations,*Doubleday*(2004)
- [Tian 15] Tian,T.,Zhu,J.,Xia,F.,Zhuang,X.and Zhang,T.:Crowd Fraud Detection in Internet Advertising,*International Conference on World Wide Web*,Vol.24,pp.1100-1110(2015)
- [Tolomei 17] Tolomei,G.,Silvestri,F.,Haines,A.and Lalmas,M.:Interpretable predictions of tree-based ensembles via actionable feature tweaking,*In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM*,pp.465-474(2017)
- [Wang 18] Wang,X.,He,X.,Feng,F.,Nie,L.and Chua,T.S.:Tem:Tree-enhanced embedding model for explainable recommendation,*In Proceedings of the 2018 World Wide Web Conference on World Wide Web*,pp.1543-1552(2018)
- [Wang 19] Wang,X.,Du,Y.,Zhang,L.,Li,X.,Zhang,M.and Dong,J.:Exploring content-

based video relevance for video click-through rate prediction, *Proceedings of the 27th ACM International Conference on Multimedia*, pp.2602–2606(2019)

[Wu 16] Wu, K., Vassileva, J., Zhao, Y., Noorian, Z., Waldner, W. and Adaji, I.: Complexity or simplicity? Designing product pictures for advertising in online marketplaces, *Journal of Retailing and Consumer Services*, Vol.28, pp.17–27(2016)

[Xia 20] Xia, B., Seshime, H., Wang, X. and Yamasaki, T.: Click-through rate prediction of online banners featuring multimodal analysis, *International Journal of Semantic Computing*, Vol.14, No.1, pp.71-91(2020)

[Zhang 17] Zhang, H., Cao, X., Ho, J.K.L. and Chow, T.W.S.: Object-level video advertising: An optimization framework, *IEEE Transactions on Industrial Informatics*, Vol.13, No.2, pp.520–531(2017)

[Zhang 18] Zhang, D., Qian, L., Mao, B., Huang, C., Huang, B. and Si, Y.: A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGboost, *IEEE Access*, Vol.6, pp.21020-21031(2018)

[馬場 15] 馬場惇, 岩崎祐貴, 杉尾樹, 北出庫介, 福嶋剛: 検索連動型広告におけるテキスト自動生成とその評価指標の検討, 人工知能学会全国大会論文集 (2015)

[電通 20] 株式会社電通グループ, 2020年日本の広告費 (2021)

[電通 21] 株式会社電通グループ: 電通グループ「世界の広告費成長率予測 (2020~2022)」改定版を公表 電通グループHP, <https://www.group.dentsu.com/jp/news/release/000509.html> (accessed 2021/11/19)

[富士通総研 21] 株式会社富士通総研: インターネット広告とは, 富士通総研HP, <https://www.fujitsu.com/jp/group/fri/report/cyber/basic/information/01.html> (accessed 2021/11/19)

- [原 18] 原聡, 説明可能 AI(Explainable AI), 人工知能学会誌, Vol.34, No.4(2018)
- [原 19] 原聡, 機械学習における解釈性 (Interpretability in Machine Learning), 人工知能学会誌, Vol.33, No.3(2019)
- [堀田 08] 堀田創, 野澤貴, 萩原将文:ニューラルネットワークを用いた位置情報に基づくインターネット広告配信システム, 日本知能情報ファジィ学会誌, Vol.20, No.3, pp.347-356(2008)
- [稲葉 19] 稲葉通将:ユーザに上手に説得される人狼知能対話システムの構築, 第 87 回言語・音声理解と対話処理研究会, pp.7-12(2019)
- [井上 02] 井上哲浩:インターネット広告の有効性に関する実証研究, 商学論究, Vol.50, No.1-2, pp.279-303(2002)
- [石井 16] 石井真人:バナー広告の印象と画像特徴の関係, 相模女子大学メディア情報研究, Vol.2, pp.1-8(2016)
- [石川 18] 石川葉子, 水上雅博, 吉野幸一郎, Sakti Sakriani, 鈴木優, 中村哲:感情表現を用いた説得対話システム, 人工知能学会論文誌, Vol.33, No.1, DSH-B_1-9(2018)
- [石川 19] 石川隆一, 和泉興, 林秀和, 福田宏幸:デザイナーが持つ審美眼のモデル化, 第 33 回人工知能学会全国大会論文集 (2019)
- [岩崎 18] 岩崎祐貴:深層学習による Facebook 広告の CTR 予測, 第 32 回人工知能学会全国大会論文集 (2018)
- [形山 01] 形山忠寛, 足田光伯:インターネット広告手法に関する一考察, 情報処理学会全国大会講演論文集, Vol.63, No.1, pp 297-298(2001)
- [株式会社電通デジタル 20] 株式会社電通デジタル:バナー広告生成ツール ADVANCED CREATIVE MAKER®2.0 は, サーバーレス移行でどこまで進化したのか? (2020), <https://www.dentsudigital.co.jp/topics/2020/0325-000411/> (accessed

2021/2/5)

[株式会社サイバーエージェント 20] 株式会社サイバーエージェント:AIで広告クリエイティブ制作を一変,報酬は広告効果がでた時のみの料金体系「極予測 AI」の提供を開始(2020),<https://www.cyberagent.co.jp/news/detail/id=24647>(accessed 2021/2/5)

[松原 20] 松原仁:『TEZUKA 2020』プロジェクト 一人間と AI が協力してマンガを描く一,人工知能学会誌,Vol.35,No.3,pp.391-394(2020)

[宮西 14] 宮西一徳,高野雅典,吉田岳彦:大規模リワード広告システムにおける行動履歴と広告属性を利用したコンバージョン予測モデルの構築,第29回人工知能学会全国大会論文集(2014)

[森下 21] 森下光之助:機械学習を解釈する技術~予測力と説明力を両立する実践テクニック,技術評論社(2021)

[中島 18] 中島ゆき:国勢調査から消えた「平成の職業」,
https://chikouken.org/report/report_cat01/8795/(2018)(accessed 2021/9/8)

[中島 19] 中島ゆき:国勢調査より「平成で新たに誕生した職業」前編,
https://chikouken.org/report/report_cat01/9065/(2019)(accessed 2021/9/8)

[中島 19] 中島ゆき:国勢調査より「平成で新たに誕生した職業」後編,
https://chikouken.org/report/report_cat01/9068/(2019)(accessed 2021/9/8)

[中根 04] 中根雅夫:インターネット広告の現状分析——実態調査を中心に——,國土館大學政経論叢,Vol.16,No.4,pp.145-169(2004)

[中岡 08] 中岡伊織,氏平淳子,串田淳一,亀井且有:「ついクリックしたくなる」バナー広告の感性評価およびそのモデル構築,日本知能情報ファジィ学会ファジィシステムシンポジウム講演論文集,Vol.24,pp.130(2008)

[中畝 98] 中畝弘,坂本啓,岸晃司:インターネット広告の経済性に関する検討,情報処理学

会全国大会講演論文集,Vol.56,pp.366-367(1998)

[西村 10] 西村洋一:インターネット利用者が検索連動型広告を回避する要因の検討:広告への態度の影響, 広告科学,Vol.52,pp.15-30(2010)

[岡田 12] 岡田美樹, 木下武志:バナー広告におけるクリック率についてのトランジションの影響:オーバーラップとワイプの効果について, 山口大学工学部研究報告 (2012)

[大坪 21] 大坪直樹, 中江俊博, 深沢祐太, 豊岡祥, 坂元哲平, 佐藤誠, 五十嵐健太, 市原大暉, 堀内新吾:XAI(説明可能な AI)-そのとき人工知能はどう考えたのか?, リックテレコム (2021)

[折原 20] 折原良平:特集「AI でよみがえる手塚治虫」, 人工知能学会誌,Vol.35,No.3,pp.391-394(2020)

[大槻 19] 大槻知明:移動通信の未来, 日本機械学会誌,Vol.122,No.1207,pp.10-11(2019)

[近江 15] 近江健太郎, 大木榮二郎:インターネット上の有害広告に対するオプトアウト方式導入の提案, 情報処理学会第 77 回全国大会講演論文集,Vol.77,No.1,pp.697-698(2015)

[李 21] 李津娥:広告を避ける心理:インターネット広告の回避をもたらす要因の検討, 慶応義塾大学メディア・コミュニケーション研究所紀要,Vol.71,pp.117-130(2021)

[坂本 17] 坂本真樹, 玉川奨:特集「広告と AI」にあたって, 人工知能学会誌,Vol.32,No.4,pp.471-472(2017)

[崎濱 19a] 崎濱栄治, 鹿島浩之:An empirical study on fake review detection by latent dirichlet allocation, 京都大学数理解析研究所講究録,Vol.2124,pp.17-27(2019)

[崎濱 19b] 崎濱栄治, 川崎泰一, 本橋永至:コンピュータービジョンによる広告画像要素のクリック訴求効果の検証, 応用統計学,Vol.48,No.3,pp.59-70(2019)

[崎濱 20] 崎濱栄治, 川崎泰一, 本橋永至:アンサンブル学習と LDA の統合による動画広告効果の要因分析, 人工知能学会論文誌,Vol.36,No.3,pp.1-8(2020)

- [佐藤 17] 佐藤敏紀, 橋本泰一, 奥村学:単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討, 言語処理学会第 23 回年次大会 (2017)
- [週刊現代 21] オックスフォード大学が認定 あと 10 年で「消える職業」「なくなる仕事」
<https://gendai.ismedia.jp/articles/-/40925>(accessed 2021/9/8)
- [田頭 13] 田頭幸浩, 山本浩司, 小野真吾, 塚本浩司, 田島玲:オンライン広告における CTR 予測モデルの素性評価, 第 11 回日本データベース学会年次大会 (2013)
- [高橋 20] 高橋あすみ, 土田毅, 末木新, 伊藤次郎:「死にたい」と検索する者の相談を促進するインターネット広告の要素は何か?, 自殺予防と危機介入, Vol.40, No.2, pp.67-74(2020)
- [田中 21] 田中絵里菜:K-POP はなぜ世界を熱くするのか, 朝日出版社 (2021)
- [谷口 17] 谷口和輝, 大田和寛, 山口光太:人工知能による新しい広告クリエイティブ, 人工知能学会誌, Vol.32, No.4, pp.501-508(2017)
- [手塚 20] 手塚眞:AI は天才を生むか～人と AI の共同創作, 人工知能学会誌, Vol.35, No.3, pp.418-421(2020)
- [戸田 10a] 戸田淳:わが国におけるインターネット広告の歴史的変遷とその本質 (その 1), 情報文化学会誌, Vol.17, No.1, pp.49-54(2010)
- [戸田 10b] 戸田淳:わが国におけるインターネット広告の歴史的変遷とその本質 (その 2), 情報文化学会誌, Vol.17, No.2, pp.26-31(2010)
- [図書印刷 21] 図書印刷株式会社:インターネット広告の種類と選び方のポイントを解説, 図書印刷 HP, <https://digital.tosho.co.jp/column/ad/2963/>(accessed 2021/11/19)
- [若江 19] 若江雅子, 森亮二, 吉井英樹:オンライン広告におけるトラッキングの現状とその法的考察—ビッグデータ時代のプライバシー問題にどう対応すべきか, 情報通信政策研究, Vol.2, No.2, pp.27-47(2019)
- [脇本 20] 脇本宏平, 川本峻頌, 張培楠:インターネット広告におけるキーワードに基づく広

告文の自動生成, 第 34 回人工知能学会全国大会論文集 (2020)

[山川 11] 山川義介: 広告クリエイティブの最適化, 日本行動計量学会第 39 回大会発表論文抄録集 (2011)

[横山 17] 横山侑政, 瀧川一学: 全部分グラフ指示子に基づく決定木勾配ブースティング, 第 31 回人工知能学会全国大会論文集 (2017)