

インターネット動画広告の視聴傾向分析のための可視化システム

三浦 梨花[†] 大矢 隼士[‡] 伊藤 貴之[†]

[†]お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

[‡] Septeni Japan 株式会社 〒160-6128 東京都新宿区西新宿 8-17-1

E-mail: [†]{g1820536, itot}@is.ocha.ac.jp, [‡]hayato.oya@septeni.co.jp

あらまし インターネット広告の市場規模は拡大を続けている。効果的なインターネット広告の制作と配信は、今後ビジネスにおいてもますます重要になるといえる。一方で、インターネット広告はターゲットを設定した上で特定の層に向けて配信するため、ターゲット設定が不適切であると広告効果が得られにくいという背景がある。そこで本研究では、インターネット動画広告データを離脱率、クリック率、コンバージョン率を用いて評価し、広告属性の組み合わせを考慮した可視化システムを開発することで、効果的なインターネット広告のターゲット設定の発見を目指す。具体例として本報告では、Yahoo!広告で配信された広告のダミーデータ 224,355 件、LINE 広告で配信された広告のダミーデータ 89,400 件を可視化した結果を報告する。

キーワード 可視化手法、動画広告、離脱率、クリック率、コンバージョン率

1. はじめに

インターネット広告の市場規模は年々拡大を続けている。電通の報告[1]によると、2021 年、インターネット広告費は 2 兆 7,052 億円に達し、マスコミ四媒体広告費の総計 2 兆 4,538 億円を初めて上回った。中でも、動画広告は前年比 132.8% の 5,128 億円と大きく伸長し、初めて 5,000 億円を突破した。このような背景から、効果的なインターネット広告の制作と配信は今後ビジネスにおいてもますます重要になる。

松本ら[2]は、マーケティングデータをクラスタリングと主成分分析を用いて分析することで、消費者の特徴を把握した。そして、広告の影響を受けた消費者がどのようなパーソナリティ及び消費価値観の特徴を有しているかについて明らかにした。その結果、広告ごとに効果的な宣伝手法が異なっていることを示唆した。また、Eric ら[3]は、クラスタ分析を用いることによって、ユーザに嫌われる広告や、反対に、ユーザに好まれる広告の特徴を発見した。以上のような研究からもいえるように、効果的な広告活動を進めるためには、それぞれの広告に適した配信手法を選択することや消費者に好まれる広告を作成することが求められる。

また、インターネット広告は、テレビなどと違い特定の層に向けて広告を配信するため、広告ごとに適切なターゲットを設定し、配信を行わないと期待した効果を得られないのが実情である。その上、インターネット広告のターゲット設定は、「女性または男性」のような二者択一の単純なターゲット設定ではなく、「女性 × 年齢」というような複数の属性を組み合わせたターゲット設定である。そのため、どの属性の組み合わせが広告効果を発揮したのかを分析結果から見つけることが難しい。一方で、広告属性の組み合わせを考慮して、広告データを可視化分析することができる手法は

現状少ない。

そこで本報告では、動画広告の離脱率、クリック率、コンバージョン率を用いて包括的に動画広告を評価した上で、広告属性の組み合わせを考慮した可視化システムを開発し、データを観察した。

本報告の主な貢献は以下の通りである。

- 以下の 3 つの指標
 - ・ 离脱率（動画再生を途中で停止する割合）
 - ・ クリック率（広告をクリックする割合）
 - ・ コンバージョン率（アプリのインストール、商品購入、契約締結などが成立する割合）
- から動画広告を包括的に評価しながらデータを観察したこと。
- 広告属性の組み合わせを考慮して、広告データの可視化システムを開発したこと。

本報告の構成は以下の通りである。2 章では関連研究について、3 章では提案手法について、4 章では実行結果と分析結果から得た考察について、5 章では本報告のまとめと今後の課題について述べる。

2. 関連研究

本章では関連研究として、広告データの可視化手法と広告効果向上を目的としたデータ活用手法について述べている論文を紹介するとともに、本研究との差異を述べる。

2.1 広告データの可視化手法

本研究で利用している動画広告データは、多次元データである。効果的な多次元データの可視化手法については、今まで様々な議論がなされてきた。また、本研究のような、広告データに特化した可視化分析の研究も少数ではあるもののいくつか存在する。Di ら[4]

の研究では、Targeting Vis というインタラクティブな分析ツールを提案した。このツールを利用するによって広告アナリストは、ターゲット広告の配信データを視覚化し、広告主の配信パターンを調査することで、広告をより効果的にする方法を発見することができる。Targeting Visにおいても、本研究と同様にターゲットの組み合わせパターンを可視化する手法が用いられており、この手法が効果的な広告設定の発見に有用であることを示唆している。Shangsong ら[5]は、マルチビューのインタラクティブシステム MulUBA を提案した。この可視化システムは、広告アナリストが、消費者のオンラインショッピング行動と属性について分析することを、複数の可視化を組み合わせることで支援している。一方でこれらの研究は、広告効果の指標として、クリック率やコンバージョン率は用いているものの、広告の離脱率に関しては言及していない。本研究においては、離脱率、クリック率、コンバージョン率の 3 つの指標から動画広告の視聴傾向を包括的に可視化することで、効果的な広告設定を発見することを目的としている。

2.2 広告効果向上を目的としたデータの活用

広告効果の向上を目指した研究は、2.1 節で述べたような可視化分析手法の研究以外にも、今までに多くなされてきた。Kitada ら[6]の研究では、マルチタスク学習を用いた不均衡データに対する学習と conditional attention 機構によって、広告クリエイティブの効果を正しく評価する枠組みを提案した。また、石川ら[7]は、Yahoo!プロモーション広告（現 Yahoo!広告）への個人の広告出稿において得られる 10 種類の基本広告データの中で、広告主に定量的な指標を明確に与えるデータについて議論した。その結果、コンバージョン予測に大きく関与する指標の組み合わせを発見した。Eric ら[3]は、どのような広告が消費者を不快にさせるのかをクラスタ分析により発見した。これにより、良い広告が集まっているクラスタには、シンプルでデザイン性に優れているものや消費者の興味に関連しているものが多いことがわかった。一方で、悪いとされる広告が集まっているクラスタには、リスト記事や政治的なものがあったことを発見した。前述した通り、今まで多様な観点から効果的な広告制作支援の研究が進められてきたが、それに対して本研究では、可視化分析のアプローチから広告効果の向上を目指す。

3. 提案手法

3.1 使用するデータ

本研究ではデータセットとして、LINE 広告と Yahoo! 広告で実際に配信された動画広告データと同様の形式

のダミーデータを利用している。

3.2 データの前処理

本研究では、大きく分けて動画広告の離脱率とクリック率、コンバージョン率の 3 つの指標を新たに適用した。本節では、それぞれの指標について詳しく述べる。

3.2.1 動画広告の離脱率

本研究では、動画広告の離脱率を用いて広告を評価する。動画広告の離脱率の傾向が似ているクラスタを生成しデータを分析するために、表 1 の 4 つの変数を用いて、以下に記載している動画広告離脱率に関する 6 つの変数 $R1 \sim R6$ を算出した。

表 1 離脱率の計算に利用する 4 つの数値

V_{p25}	動画広告の 25% 時点まで再生された回数
V_{p50}	動画広告の 50% 時点まで再生された回数
V_{p75}	動画広告の 75% 時点まで再生された回数
V_{p95}	動画広告の 95% 時点まで再生された回数

$$R1 = \frac{V_{p25} - V_{p50}}{V_{p25}}$$

$$R2 = \frac{V_{p25} - V_{p75}}{V_{p25}}$$

$$R3 = \frac{V_{p25} - V_{p95}}{V_{p25}}$$

$$R4 = \frac{V_{p50} - V_{p75}}{V_{p25}}$$

$$R5 = \frac{V_{p50} - V_{p95}}{V_{p25}}$$

$$R6 = \frac{V_{p75} - V_{p95}}{V_{p25}}$$

3.2.2 クリック率 / コンバージョン率

動画広告の離脱要因は、ポジティブな要因とネガティブな要因の 2 つに分けられる。ネガティブな離脱とは、広告効果を發揮せず、単に動画広告を離脱されてしまったことを指す。一方で、ポジティブな離脱とは、広告がクリックされたことによって起こる広告離脱を指す。広告動画の離脱率からのみ広告を評価した場合、

その離脱がポジティブな離脱であるかネガティブな離脱であるか区別することができない。一般的にインターネット広告は、広告の表示数に対するクリック数であるクリック率(CTR)や広告をクリックした後に広告配信の目的が達成される率であるコンバージョン率(CVR)で評価されることが多い。そこで、動画広告の離脱率、クリック率、コンバージョン率を合わせて分析することで、ポジティブな離脱とネガティブな離脱を区別することが可能であると考えた。よって本研究では、動画広告の離脱率、クリック率、コンバージョン率を用いて包括的に広告を評価する。

3.3 次元削減・クラスタリング

3.2.1 節で算出した動画広告の離脱率に関する 6 つの変数 $R1 \sim R6$ を次元削減し、6 次元から 2 次元に圧縮する。また、 $R1 \sim R6$ の 6 つの変数にクラスタリングを適用することで、同一クラスタに離脱傾向が似た広告が属するようになっている。これにより、本研究では、動画広告の離脱率の傾向が似ている広告の中で、それぞれどのような特徴が見られるかについて観察する。また、次元削減手法として主成分分析を、クラスタリング手法として k-means 法を用いている。

3.4 可視化

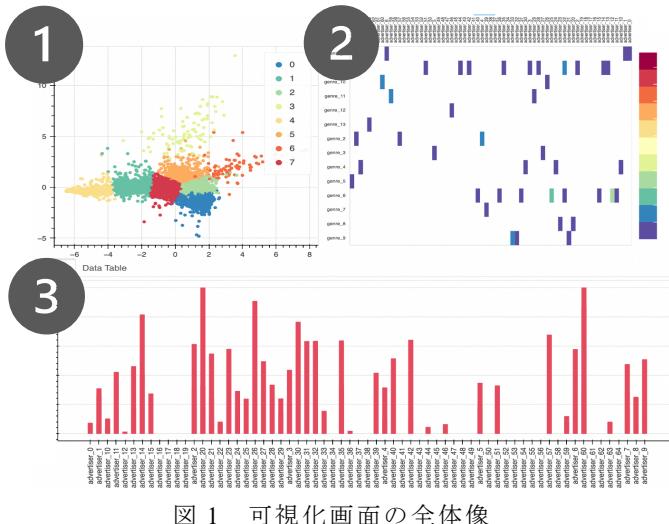


図 1 可視化画面の全体像

row index	advertiser	genre	sub_genre	web_app	conversions	video_views	ctr	cvr	impressions	costs	clicks	row index
5	advertiser_0	genre_0	sub_genre_0	1.0	214.0	1.1	7.2	4209.5	6708.5	29.5	2	
5	advertiser_1	genre_0	sub_genre_0	1.0	1510.7	0.1	6.2	18263.7	21172.3	16.7	9	
5	advertiser_1	genre_1	sub_genre_1	1.6	832.9	0.6	26.4	4853.4	3993.8	27.6	21	
5	advertiser_1	genre_0	sub_genre_1	1.5	513.4	0.7	30.8	3215.5	2442.9	11.0	22	
5	advertiser_1C	genre_4	sub_genre_8	1.1	11647.9	0.0	3.3	99186.4	37651.9	45.5	31	
5	advertiser_1C	genre_4	sub_genre_8	1.6	2189.2	0.1	10.1	22270.3	24466.6	22.2	32	
5	advertiser_1C	genre_4	sub_genre_8	3.0	4445.7	0.3	3.9	43442.1	78217.6	105.3	33	
5	advertiser_11	genre_6	sub_genre_9	1.0	600.9	0.9	34.1	9991.3	1323.1	49.5	44	
5	advertiser_11	genre_6	sub_genre_9	1.0	339.7	1.1	42.1	4801.7	492.4	28.1	45	

図 2 図 1 の③において、タブを切り替えて表示したデータテーブル

本研究では、Python の可視化ライブラリである

Bokeh を用いて可視化システムを開発した。Python は 3.9.6, Bokeh は 2.4.3 を使用した。図 1 が本研究の可視化の全体像である。また、図 1 の下部の可視化画面でタブを切り替えることによって、図 2 のようなデータテーブルを表示させることができる。①～④の 4 つの可視化について以下に述べる。

- ① 3.3 節で述べた手法で、広告データに次元削減・クラスタリングを適用し、散布図で可視化している。動画広告の離脱傾向が似ている広告データが同一のクラスタに属しており、クラスタごとに固有の色が割り当てられている。
- ② 可視化①でプロットされているクラスタのうち 1 つを選び、ヒートマップでそのクラスタに属している広告データのクリック率を可視化している。図 1 の可視化画面においては、縦軸が広告ジャンルの属性で、横軸が広告主の属性である。図 3 に示すように、クリック率が高い属性の組み合わせには、明るい色が、クリック率の低い属性の組み合わせには暗い色が割り当てられている。
- ③ 広告主ごとのコンバージョン率を棒グラフで可視化した結果。可視化②のヒートマップでデータを選択することで、棒グラフにて選択されたデータがハイライトされる。
- ④ 可視化②で表示されたデータの詳細をデータテーブルで確認することができる。

4. 実行結果・考察

本章では、実行結果の一部と実行結果から得られる知見について述べる。例として、本報告では、Yahoo!広告で配信された動画広告データのダミーデータ 224,355 件、LINE 広告で配信された動画広告データのダミーデータ 89,400 件を可視化分析した結果を紹介する。

4.1 Yahoo!広告で配信された動画広告の可視化結果

図 4 は、Yahoo!広告で配信された動画広告データのダミーデータから、3.2.1 節の手法で算出した動画離脱率に関する 6 つの変数 $R1 \sim R6$ を算出し、主成分分析・クラスタリングを適用し、散布図で可視化した結果である。また、表 2 に第一主成分(PC1)と第二主成分(PC2)

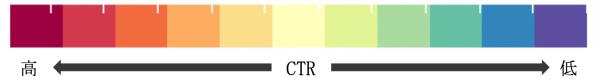


図 3 ヒートマップで用いたカラーマップ

の固有ベクトルを算出した結果を示す。表 2 に示した。PC1 は $\mathcal{R}1 \sim \mathcal{R}3$ の絶対値が高く、PC2 は相対的に $\mathcal{R}4 \sim \mathcal{R}6$ の絶対値が高いので、PC1 は動画全体における離脱率、PC2 は後半における離脱率といえる。図 4 の第一主成分において、左にいくほど、動画広告の離脱率が高い広告データが、また、右にいくほど動画広告の離脱率が低い広告データがプロットされている。

図 5 は、図 4 の散布図で可視化された 8 つのクラスタのうち、動画広告の離脱率が最も小さいとされるクラスタ 1 に属する広告のクリック率をヒートマップで可視化したものである。縦軸は動画広告のジャンルで、横軸は広告主である。この結果から動画広告離脱率の低い広告の中でも、クリック率に差が生じていることがわかる。具体的には、クリック率の高い広告は、「genre_8」かつ「advertiser_12」や、「genre_4」かつ「advertiser_4」の広告が該当する。

一方で、動画広告の離脱率が高いとされるクラスタ 3 を選択して、ヒートマップで可視化した結果を図 6 に示す。広告の離脱率が高い広告データの中にも、「genre_13」かつ「advertiser_48」のようなクリック率が高い属性の組み合わせが存在することがわかる。このような属性の組み合わせの広告は、クリックされたことによって離脱が起きた回数が多いことを示唆している。よって、これらの広告属性の組み合わせの広告は、クリックにつながっており、効果的な広告活動であるといえる。

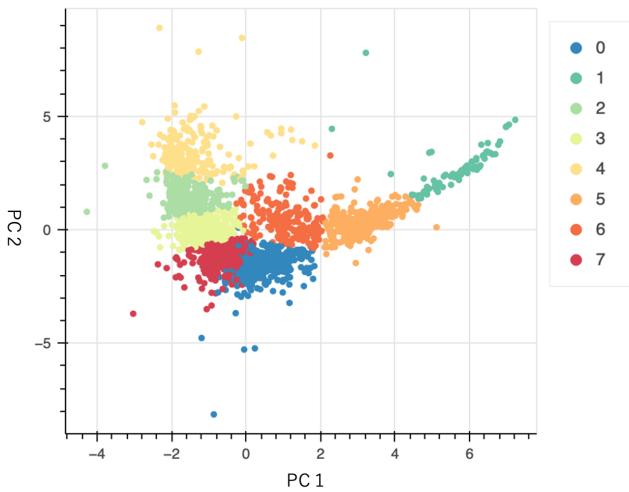


図 4 Yahoo!広告で配信された広告データと同様の形式のダミーデータに主成分分析 / クラスタリングを適用して散布図で可視化した結果

表 2 Yahoo!広告で配信された広告データと同様の形式のダミーデータを主成分分析した際の固有ベクトルの値

	PC1	PC2
$\mathcal{R}1$	-0.511	0.241
$\mathcal{R}2$	-0.554	0.083
$\mathcal{R}3$	-0.555	-0.008
$\mathcal{R}4$	-0.287	-0.504
$\mathcal{R}5$	-0.153	-0.613
$\mathcal{R}6$	0.134	-0.553

図 6 で可視化されたデータのうち、クリック率が高い広告属性の組み合わせを指定して選択し、コンバージョン率を可視化した棒グラフ上で該当データをハイライトした結果が図 7 である。このターゲット属性の組み合わせは、クリック率が大きく、コンバージョンにも結びついていることがこの可視化結果から推測される。このように本研究の可視化結果を見ることによって、広告効果を動画広告の離脱率、クリック率、コンバージョン率の観点から包括的に分析することが可能である。

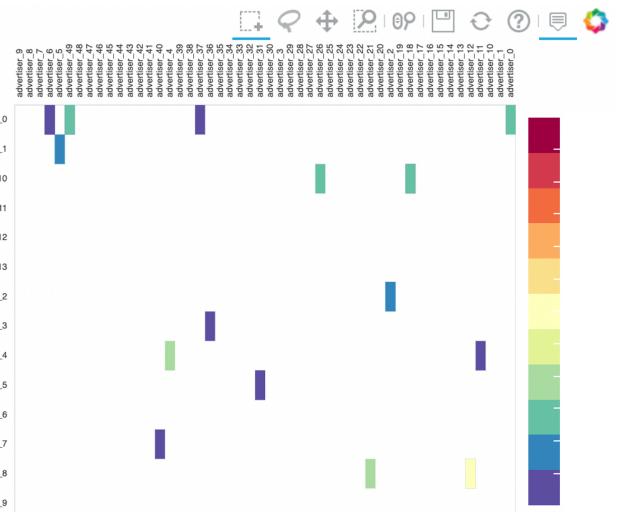


図 5 図 4においてクラスタ 1 に属する広告データのクリック率をヒートマップで可視化した結果（縦軸は広告ジャンル、横軸は広告主）

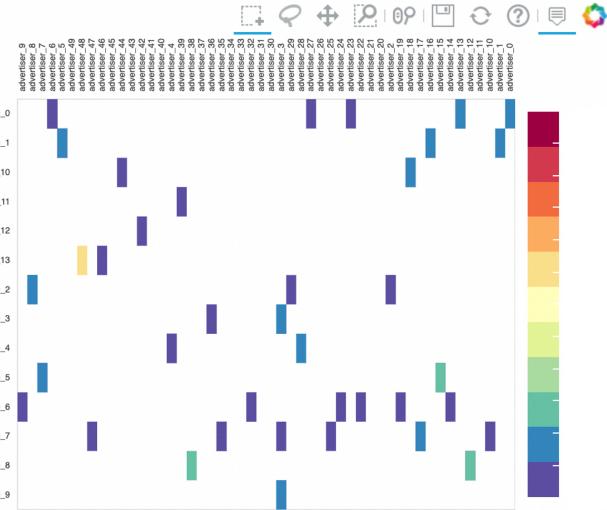


図 6 図 4においてクラスタ 3に属する広告データのクリック率をヒートマップで可視化した結果（縦軸は広告ジャンル、横軸は広告主）

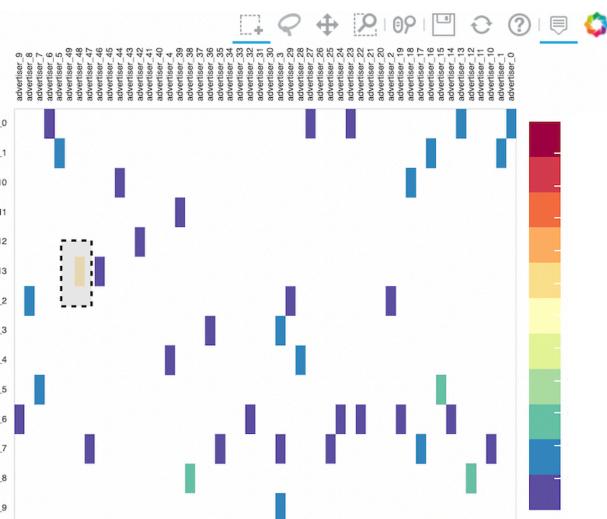


図 7 離脱率の高いクラスタ 3に属するデータのうち、クリック率の高い属性の組み合わせを選択する

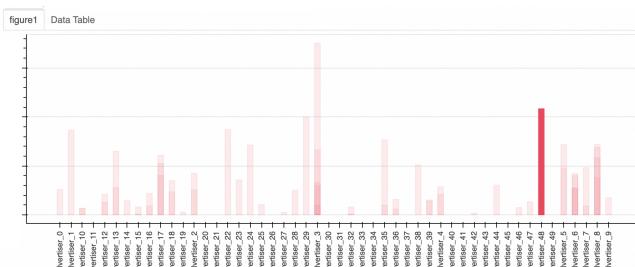


図 8 図 7において選択したデータの広告主ごとのコンバージョン率をハイライトした可視化結果

4.2 LINE 広告で配信された動画広告の可視化結果

LINE 広告で配信された動画広告データと同様の形式のダミーデータに対して、3.3 節で示した手法で可視化した結果を図 9 に示す。また、表 3 に主成分分析の固有ベクトルの結果を示す。PC1 は $\mathcal{R}1 \sim \mathcal{R}5$ の絶対値が高く、PC2 は相対的に $\mathcal{R}6$ の絶対値が高いので、PC1 は動画全体における離脱率、PC2 は後半における離脱率といえる。図 9において、左にいくほど、動画広告の離脱率が低い広告データがプロットされており、右にいくほど、動画広告の離脱率が高い広告がプロットされている。離脱率が最も低いとされるクラスタ 4 に属する広告のクリック率をヒートマップで可視化した結果が図 10 である。縦軸は広告ジャンルで、横軸は広告主である。可視化結果より、LINE 広告で配信されるデータは、比較的クリック率の高い広告が多いことがわかる。

ヒートマップの縦軸と横軸の属性を変更して可視化した結果を図 11 に示す。図 11において、縦軸は広告のサブジャンルの属性、横軸はターゲティングタイプの属性である。ターゲティングタイプの属性の説明は、表 4 に示す。可視化結果より、最もクリック率が高いのは、「sub_genre_4」と「Look Alike」の広告属性の組み合わせであることがわかる。このように本研究の可視化手法では、ヒートマップの縦軸・横軸の属性を変えることで、2 つの属性の組み合わせごとにクリック率の高い広告設定を発見することが可能である。

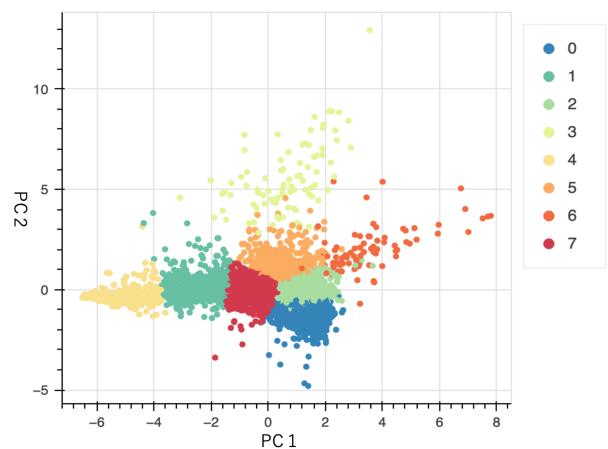


図 9 LINE 広告で配信された広告データと同様の形式のダミーデータに主成分分析 / クラスタリングを適用して散布図で可視化した結果

表3 LINE広告で配信された広告データと
同様の形式のダミーデータを主成分分析した際の
固有ベクトルの値

	PC1	PC2
R1	0.408549	-0.428766
R2	0.475229	-0.294864
R3	0.495529	-0.114966
R4	0.402646	0.136094
R5	0.409860	0.477465
R6	0.177725	0.685203

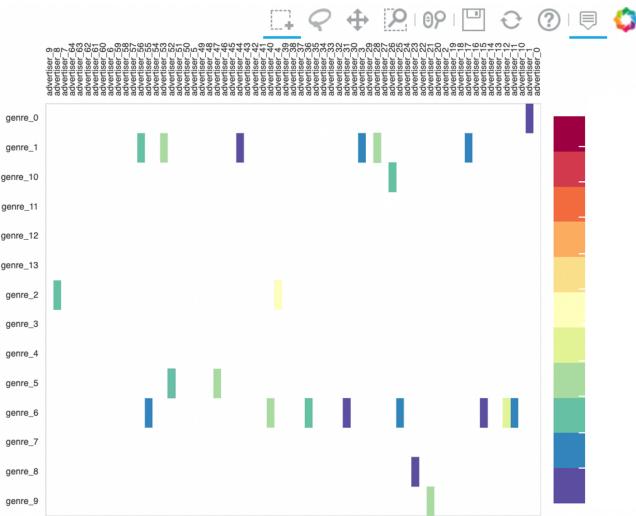


図10 図9中のクラスター4に属する広告データのク
リック率をヒートマップで可視化した結果
(縦軸は広告ジャンル, 横軸は広告主)

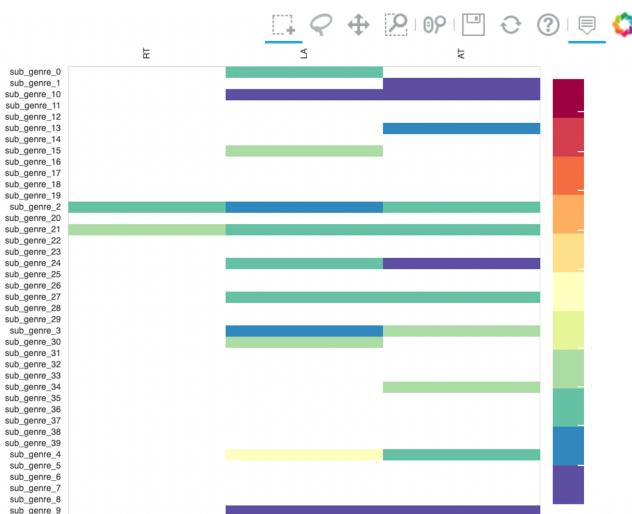


図11 図9中のクラスター4に属する広告データの
クリック率をヒートマップで可視化した結果
(縦軸は広告のサブジャンル, 横軸はターゲティングタイプ)

表4 ターゲティングタイプの説明

RT (Re Targeting)	一度見た商品などを表示する広告
LA (Look Alike)	消費者が過去に見たものや購入したものと似ているものを表示する広告
AT (Audience Targeting)	上記2つ以外の広告

5. 結論・今後の課題

本報告では、動画広告の離脱率、クリック率、コンバージョン率を合わせて包括的に広告を評価した。また、広告属性の組み合わせを考慮した可視化を実行し、その結果例を示した。その結果、以下のことがわかった。

- 離脱率が高い動画広告の中にも、クリック率の高い動画広告があること。このような動画広告は、クリックされたことによる離脱が生じている可能性があること。
- 配信するプラットフォームごとにクリック率の高い広告属性の組み合わせが異なること。

今後の課題を3点あげる。1点目は、本研究のデータに最適な次元削減手法・クラスタリング手法の選出である。現在、次元削減手法として主成分分析を、クラスタリング手法としてk-means法を用いている。しかし、文献[8]でも指摘されているように、どの次元削減手法・クラスタリング手法を適用するかによって、可視化結果は大きく変化する。そこで、現在用いている手法が適切であるかを評価する必要がある。今後はより多様な手法を比較し、本研究データに適切な手法を模索したいと考えている。2点目は、可視化ツールのインタラクティブ性の向上である。現状の可視化ツールでは、データの選択をインタラクティブにすることはできない。しかし、今後はユーザが自由に可視化するデータの選択をできるように変更し、より効果的な可視化分析が可能になるようにならうと考えている。3点目は、可視化③への平行座標プロットの適用である。平行座標プロットを用いることで、多次元データの情報を損なわずに、データ間の関係を解釈することができ、より詳細なデータ分析が可能となると考えている。

参考文献

- [1] 電通：「2021年日本の広告費 インターネット広告 媒体費 詳細分析」(2022), <https://www.dentsu.co.jp/news/item-cms/2022007-0309.pdf>. Accessed: 2022-11-01
- [2] 松本有加, 堂前力真, 大槻明, “x-means クラスタリング及び主成分分析を用いた消費者の特徴分

析”, 情報知識学会誌, pages 200-207, 2018.

- [3] Eric Zeng, Tadayoshi Kohno, Franziska Roesner, “What makes a “Bad” Ad? User Perceptions of Problematic Online Advertising”, CHI Conference on Human Factors in Computing Systems, pages 1-24, 2021.
- [4] Di Peng, Wei Tian, Min Zhu, Yukun Ren, Xiaojian Lin, Mingzhao Li, “TargetingVis: visual exploration and analysis of targeted advertising data”, Journal of Visualization, vol.23(6), pages 1113-1127, 2020.
- [5] Shangsong Liu, Di Peng, Haotian Zhu, Xiaolin Wen, Xinyi Zhang, Zhenghao Zhou, Min Zhu, “MulUBA: multi-level visual analytics of user behaviors for improving online shopping advertising”, Journal of Visualization, vol.24(6), pages 1287-1301, 2021.
- [6] Shunsuke Kitada, Hitoshi Iyatomi, Yoshifumi Seki, “Conversion Prediction Using Multi-task Conditional Attention Networks to Support the Creation of Effective Ad Creatives”, In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD ’19, pages 2069–2077, New York, NY, USA, 2019.
- [7] 石川善一郎, 奥牧人, 河野崇, “Web リスティング広告における基本広告データを用いたコンバージョン分析と予測”, DEIM Forum, 2017.
- [8] Sylvain M. Lespinats, “CheckViz: Sanity Check and Topological Clues for Linear and Non-Linear Mappings”, Computer Graphics Forum, pages 113-125, 2011.