

# インターネット動画広告の可視化システム： 離脱傾向を踏まえた広告データの可視化分析

理学専攻 情報科学コース 2240671 三浦 梨花 (指導教員：伊藤 貴之)

## 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、インターネット広告費は 2022 年に、前年比の 114.3% である 3 兆 912 億円に達した。特に、動画広告においては前年比 115.4% の 5,920 億円と大きく伸長し、2023 年は引き続き高い成長率を保ち、6,852 億円まで増加すると予測されている。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といえる。そこで本研究では、インターネット動画広告データを対象に、下記の T1 から T3 の 3 つの可視化タスクを満たすようなインタラクティブな可視化システムを開発する。

T1: 動画広告の離脱率ごとにクリック率・コンバージョン率にどのような特徴が現れるのかを分析できるようにすること。

T2: クリック率が大きく変動するようなパラメータの組み合わせを見つけやすくすること。

T3: 見る価値のある広告データを探索しやすくすること。この可視化システムにより、効果的な配信設定の発見を目指す。具体例として Yahoo! 広告で配信された広告のダミーデータ 224,355 件を可視化した結果と、その考察を報告する。

## 2 処理手順

### 2.1 データの前処理

#### 2.1.1 動画広告の離脱率

本研究では、動画広告の離脱率の傾向が似ているクラスタを生成するために、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  $R_1 \sim R_6$  を算出した。ここで  $V_{px}$  は、動画広告の  $x\%$  時点まで再生された回数とする。

$$\begin{aligned} R_1 &= \frac{V_{p25} - V_{p50}}{V_{p25}}, R_2 = \frac{V_{p25} - V_{p75}}{V_{p25}}, \\ R_3 &= \frac{V_{p25} - V_{p95}}{V_{p25}}, R_4 = \frac{V_{p50} - V_{p75}}{V_{p25}}, \\ R_5 &= \frac{V_{p50} - V_{p95}}{V_{p25}}, R_6 = \frac{V_{p75} - V_{p95}}{V_{p25}}, \end{aligned} \quad (1)$$

#### 2.1.2 クラスタリング

2.1.1 節で算出した動画広告の離脱率に関する 6 つの変数  $R_1 \sim R_6$  に対して、k-means クラスタリングを適用する。その結果、動画の離脱傾向が似ている広告が同一クラスタに属するようになる。

#### 2.1.3 クリック率とコンバージョン率

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数であるコンバージョン率 (CVR) で評価されることが多い。本研究もこれらの指標を用いて広告を評価する。

### 2.2 可視化手法

本論文の可視化システムを図 1 に示す。可視化システムはコントロールパネルと可視化領域の 2 つの要素で構成されている。

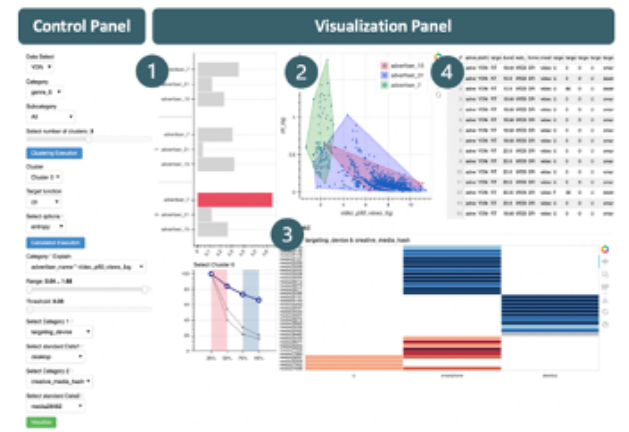


図 1: 可視化結果の全体像

#### 2.2.1 コントロールパネル

図 1 の左側に位置するコントロールパネルにて、可視化するデータの条件を設定する。具体的には、「可視化データ選択部分」、「クラスタリング実行部分」、「目的変数選択部分」、「説明変数・カテゴリ変数選択部分」、「目的変数 (CTR・CVR) の閾値変更部分」、「ヒートマップデータ選択部分」、「可視化実行部分」を有する。また、「説明変数・カテゴリ変数選択部分」では、伊藤ら [2] が提案する correlation・skinny・clumpy・entropy の 4 つの指標を用いて、可視化する価値のある説明変数とカテゴリ変数を特定する。そして、算出された 4 つの指標のスコアをソートしたものをユーザーに提示する。

#### 2.2.2 可視化領域

図 1 の右側に位置する①から④の 4 つの可視化の詳細を以下に述べる。

##### 可視化①

2 章で述べた手法を用いて、広告データにクラスタリ

ングを適用した結果を、棒グラフと折れ線グラフで可視化する。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いる。また、目的変数の値が、各ジャンルの広告の目的変数の平均値よりも高い広告は、棒グラフにおいてピンク色でハイライトされる。

### 可視化②

可視化①でプロットされているクラスタのうち1つを選択し、散布図でデータを可視化する。縦軸には目的変数（CTR または CVR）を、横軸には、コントロールパネル部分で選択された説明変数を設置している。また、選択されたカテゴリ変数のラベルごとに色を割り当て、包括領域を表示する。また、可視化に用いる説明変数とカテゴリ変数の選択肢は、2.2.1 節で述べた手法を用いて、説明変数とカテゴリ変数をユーザー側に提示し、見る価値のある散布図を生成することをフォローする。

### 可視化③

広告属性の組み合わせごとに広告データの CTR もしくは CVR を可視化する。ヒートマップの縦軸と横軸には、コントロールパネル部分で選択された2つのカテゴリ変数をそれぞれ設定している。そして、ユーザーによって選択された広告属性の組み合わせを基準値とし、グレーで表示する。基準値よりも目的変数が低い属性の組み合わせには青色を、高い組み合わせには赤色が割り当てられている。また、コントロールパネル部分で目的変数の閾値を選択することで、表示するデータを絞り込む。閾値を上げることによって、基準値から大きくクリック率またはコンバージョン率が上昇・下降する広告データのみを表示することができる。

### 可視化④

可視化②において、可視化されたデータの詳細をデータテーブルで表示する。コントロールパネル部分の Range スライダーで値の範囲を絞ることで、可視化②における目的変数が高い値のデータや低い値のデータのみを確認することができる。

## 3 実行例

図 2 は、Yahoo!広告で配信された広告のダミーデータ 224,355 件を 3 章で示した手法で可視化した結果である。図 2 の可視化①の結果から、選択されたクラスタ 0 は、離脱傾向が緩やかな広告群であることがわかる。また、クラスタ 0 に属する広告主 7 の広告は、クリック率が高い傾向にある。次に図 2②の可視化結果を観察する。広告主 31 は散布図の左下に点群があることから、動画広告の 25%時点から 50%時点までの離脱率は低いが、クリック率も低い。一方で、広告主 7 の広告は、クリック率が他の 2 つの広告主の広告よりも高いことがわかる。最後に図 2③を観察する。図 2③は、縦軸が広告の種類、横軸が広告のターゲットデバイスで CTR を可視化したヒートマップである。ヒートマップの基準値は「広告種類：27939」と「ターゲットデバイス：U(U：未設定)」の組み合わせの広告に設定している。可視化結果から、「広告種類：27939」の広告は「ターゲットデバイス：U(U：未設定)」で配信するよりも「ターゲットデバイス：smartphone」で配信した方が CTR が高いことから、smartphone に向けた配信が効果的であることがわかる。このように可視

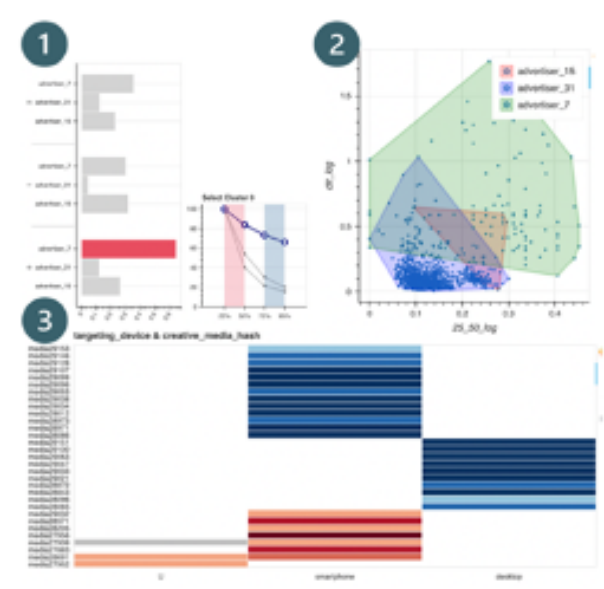


図 2: 可視化結果の全体像

化①から可視化③を組み合わせることで、離脱傾向を出発点として、どのような属性の広告がクリックされやすいのかを発見することができる。

## 4 まとめ

本研究では、T1~T3 の 3 つの可視化タスクに沿って、インターネット動画広告データを対象にした可視化システムを開発した。インタラクティブにデータを探索して、複数の可視化を用いることによって、広告属性ごとの CTR や CVR の傾向を観察することができた。今後の課題を 2 つ述べる。1 つ目は、可視化②部分で絞り込んだデータの特徴を一目で把握できる可視化を可視化④で実現することである。2 つ目は、動画広告の 25%時点までの離脱が大きなデータを見つけやすくすることである。

謝辞：本研究にあたり研究に対する多くのアドバイスをいただいた、共同研究者である Septeni Japan 株式会社大矢氏に、感謝いたします。

## 参考文献

- [1] 電通：「2022 年 日本の広告費 インターネット広告媒体費 詳細分析」(2023), <https://www.dentsu.co.jp/news/item-cms/2023007-0314-2.pdf>. (Accessed : 2023-12-21).
- [2] Takayuki Itoh, Asuka Nakabayashi, Mariko Hagita, “Multidimensional Data Visualization Applying a Variety-Oriented Scatterplot Selection Technique”, Journal of Visualization, Vol. 26, No. 1, pages 199-210, 2023.