

CMの広告効果に関する 要因の探究と予測モデルの考案

同志社大学文化情報学部文化情報学科
内藤宏明





CONTENTS

01 研究背景

媒体別広告費の実態 / データ概要 /

NRIの広告効果予測モデル / 研究目的

02 データ加工・基礎集計

データ加工 / CMの広告効果 /

セミパラメトリックなDID推定量 / 広告効果の基礎集計

03 本分析①

CMの特徴量を用いた広告効果の要因分析

04 本分析②

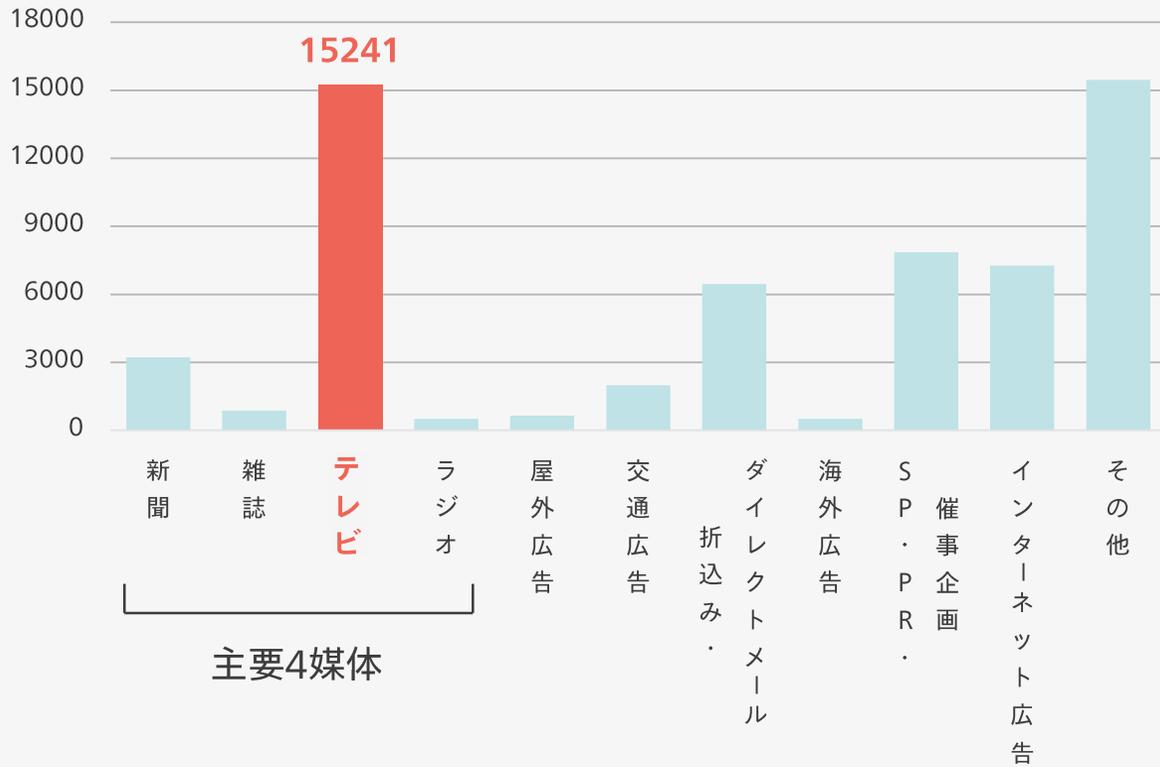
広告効果予測モデルの考案

05 まとめ

CM制作の指標提言 / CMの具体例 / 予測モデルの活用

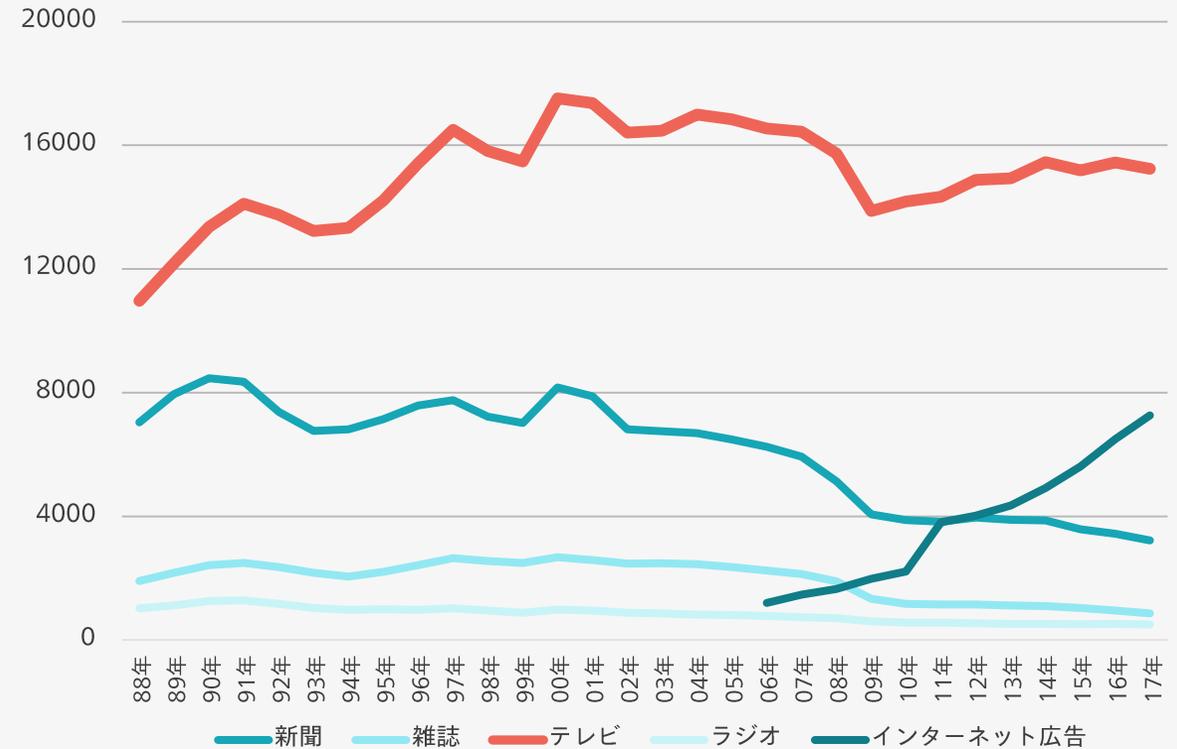
媒体別広告費の実態 (経済産業省調べ[1])

2017年の年次広告費 (単位：億円)



→ 「 広告媒体別に見たとき、近年最も広告費が大きいのはテレビである。」

1988 - 2017年の媒体別広告費 (主要4媒体 + インターネット, 単位：億円)



→ 「 どの年代もテレビの広告費は他の媒体を上回り、堅調に推移している。」

経済産業省の調査結果から、**広告媒体としてのテレビに対する信頼は非常に厚く、その傾向は今後も続くと考えられる。**

広告主がCMに多額のコストを掛けるのは**テレビには広告効果 (※) がある**という考えが経験知として業界に浸透しているからである。

(※) 広告効果については後述。

テレビにおける広告とは、一般的にテレビCM (以下、CM) である。そこで、**今後も広告主はCMを放映するために多額の広告費を投じることが推察される。**

しかし、近年テレビ離れの増加が叫ばれており、**広告効果の高いCMがどんなものであるかこれまでより精緻に把握する必要に迫られる時代になってきた。**

データ概要

- | CMの広告効果を測ることができる可能性を持つデータとして野村総合研究所(以下, NRI)が主催するマーケティングコンテストで提供されるデータがある.
- | 大きく分けてアンケートデータと出稿データの2種類がある.



アンケートデータ

3000サンプルを対象としたシングルソース形式
(調査対象: 関東1都6県の男女20~59歳).

個人属性, チャンネル利用, 商品別の購入実態・購入意向などが分かる.



出稿データ

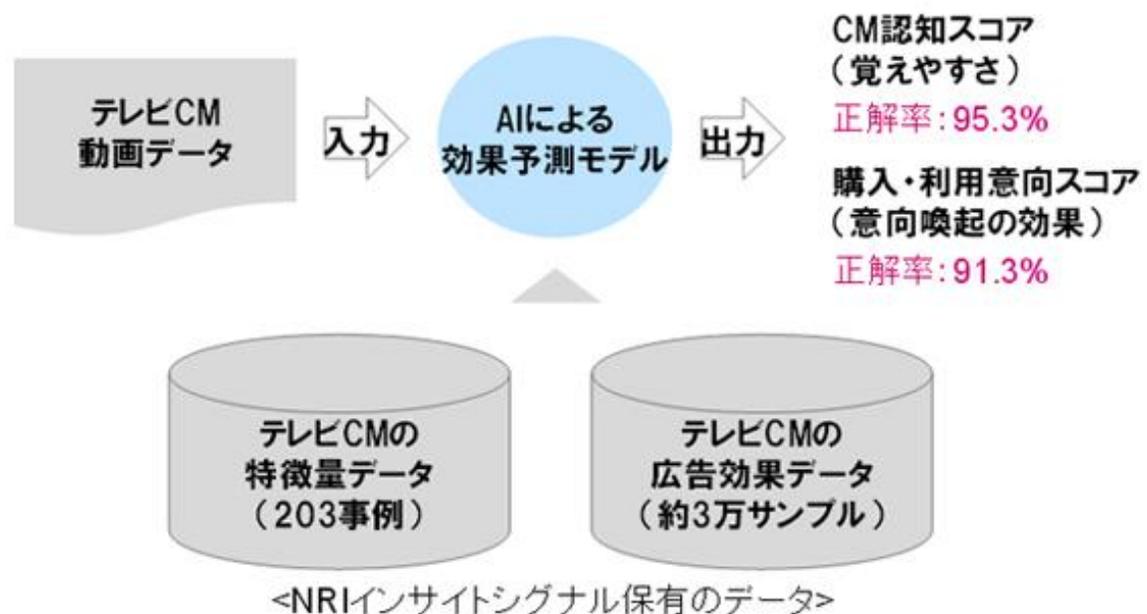
CMと雑誌広告についてどの番組(雑誌)にどの商品のCM(広告)があったのかを整理したもの.

アンケートデータと合わせて調査対象がどの番組(雑誌)を視聴して(読んで)いたのかが分かる.

NRIの広告効果予測モデル

関連研究として、2018年2月にNRIが未公開のCMの広告効果を高精度に予測するモデルを開発したことを発表している[2].

図表 テレビCMの効果予測モデルの全体像



モデルの概要

説明変数

色彩, タレント表示時間, 商品表示時間, BGM, カット数などのCMの特微量データ.

目的変数1: CM認知スコア

あるCMに10回接触したときに消費者がそのCMを知っていると回答する割合をスコア化したもの.

目的変数2: 購入・利用意向スコア

CMと接触することによって, 消費者がそのCMで扱っている商品等を買いたくなる・利用したくなる効果をスコア化したもの.

以上の背景を考慮し、次の2つの研究目的を定めた。

1 広告効果がある(ない)CMの特徴を明らかにする。

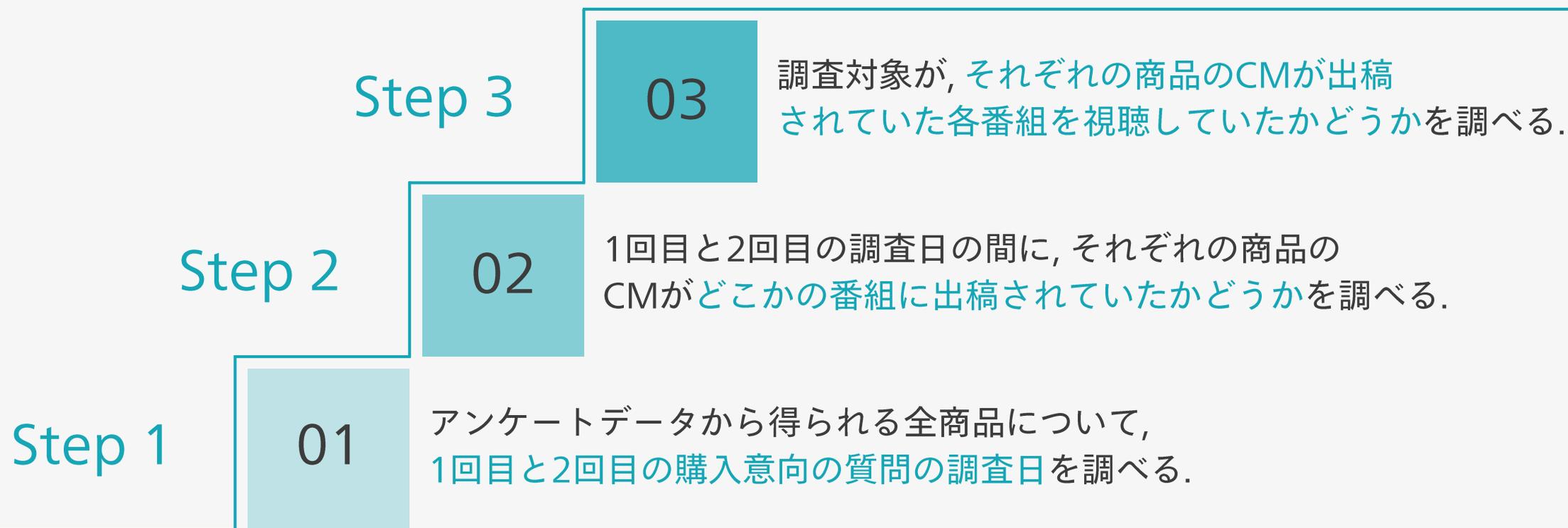
4頁でも述べたように、広告効果のあるCMの特徴を捉えることは昨今の重要なビジネス課題である。よって本研究では、CMを作る際に留意すべきことを提言する。また、前頁の予測モデルにおいて、CM認知スコアや購入・利用意向スコアを改善するためにどの説明変数をどのように操作すべきか把握するには試行を重ねる必要がある。そのため予測モデルの使用性を高めるという観点からも有意義な試みだと考えられる。

2 未公開のCMの広告効果の有無を判別するモデルを考案する。

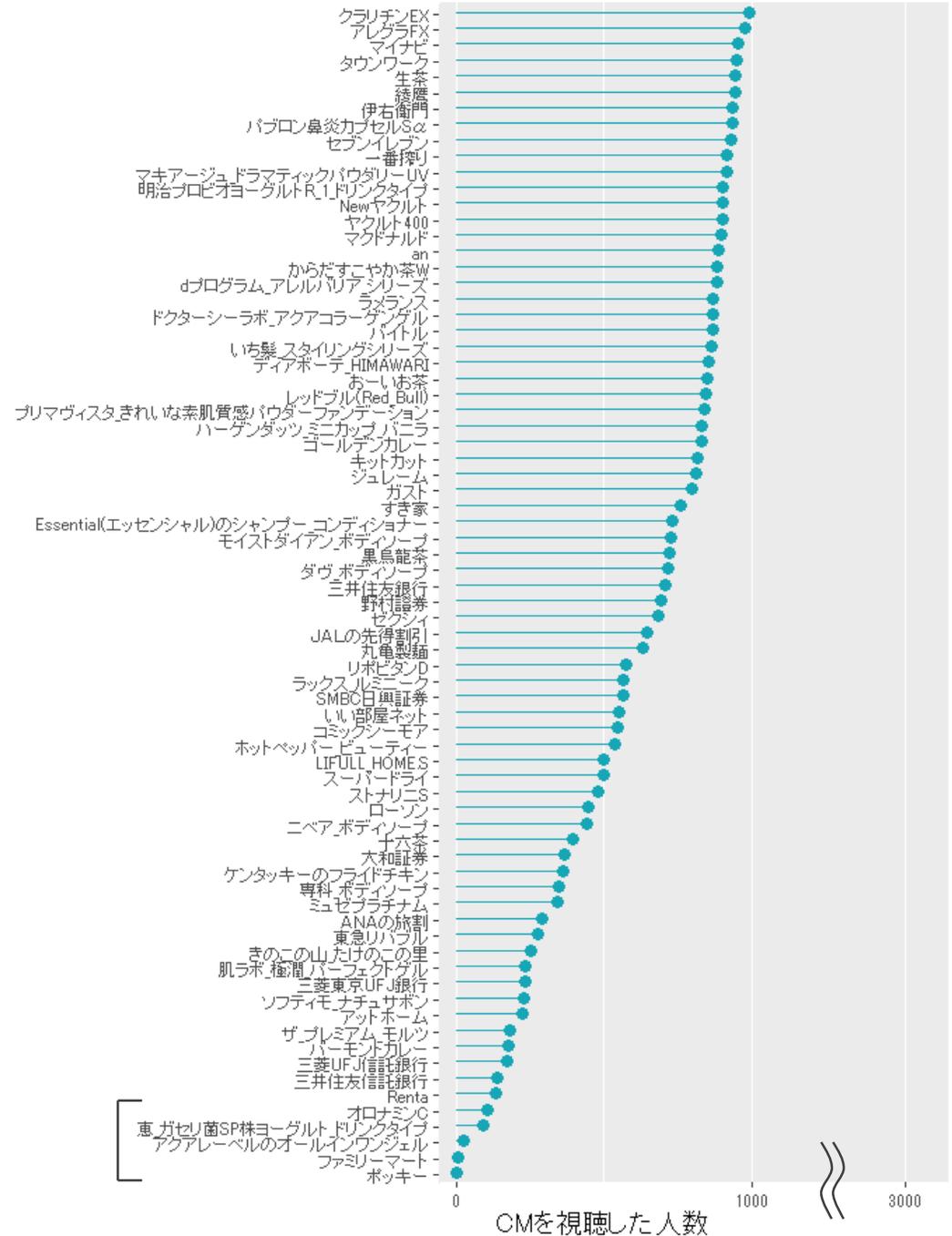
モデルを考案するモチベーションはNRIの研究と同様で、公開する前の段階で当該CMに広告効果があるかどうかを検証することである。5頁で述べたアンケートデータの中の購入意向の質問を用いて、購入・利用意向スコアに類似した広告効果指標を作成し、それを目的変数としたモデルを作成する。本研究では特に、色彩情報の抽出方法を工夫し、この情報を説明変数とした解釈可能なモデルを考案する。

データ加工のプロセス

まず各商品のCMの広告効果を求める準備として、以下の手順でデータ加工を行う。



1回目と2回目の調査の間にCMを出稿していた商品



データ加工の結果

- アンケートデータから得られる全153商品のうち、1回目と2回目の調査日の間にCMを出稿していた商品は74商品であった。
- また、CMを視聴した人数が100人未満であった以下の4商品は、**広告効果が正確に求められない恐れがあるため**以降の分析対象から外すこととした。

恵ガセリ菌SP株ヨーグルト_ドリンクタイプ、
アクアレーベルのオールインワンジェル、
ファミリーマート、ポッキー

CMの広告効果

次に, アンケートデータの購入意向の質問を用いて広告効果を定義する. 質問内容は以下の通り.

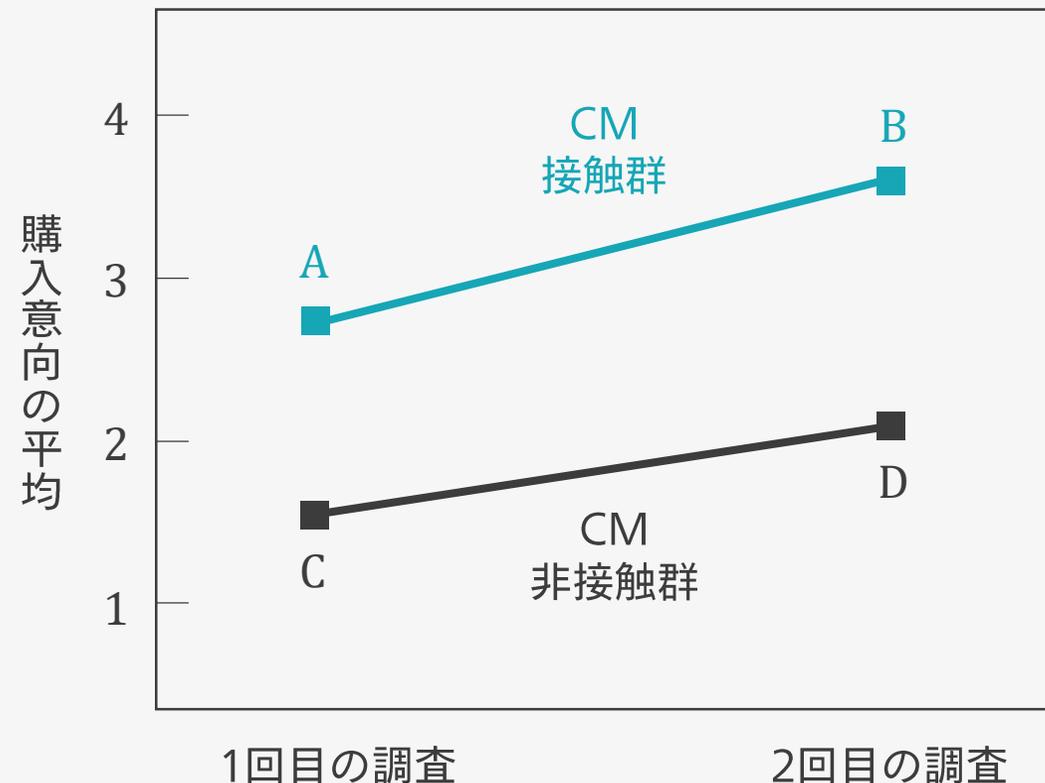
あなたは今後●●を買いたいと思いますか

1. 買いたくない
 2. わからない
 3. 買いたい
 4. ぜひ買いたい
- (※)

(※) 以後の分析のために選択肢の並びを元のデータと逆にしている.



この質問を同一の調査対象に期間を空けて2回行っている. まず, 1回目の調査と2回目の調査の間に当該商品のCMに接触したか否かで購入意向がどのように変化したのかを調べる. そこで, 調査対象を2群に分け, 次のような図の A, B, C, D を考える.



このように, 2群の1回目と2回目の調査の平均の差を比較する手法はDID分析といい, $(B - A) - (D - C)$ の値をDID推定量と呼ぶ[3].

セミパラメトリックなDID推定量

アンケートデータには調査対象の個人属性も含まれる。そこで本研究では、それらを共変量として用いることで得られる一般のDID分析よりも緩い仮定を置くセミパラメトリックなDID推定量を広告効果量とする。

	CM接触群 $z = 1$	CM非接触群 $z = 0$
y_{1b}	$z = 1$ の 事後測定 of データ	欠測
y_{0b}	欠測	$z = 0$ の 事後測定 of データ
y_a	事前測定 of データ	
x	共変量	

左図において

- a : 1回目の調査を表す指標
- b : 2回目の調査を表す指標
- y : 購入意向

とおけば、セミパラメトリックなDID推定量は

$$\frac{1}{N} \sum_{i=1}^N \frac{y_{bi} - y_{ai} z_i - e_i}{p(z = 1) (1 - e_i)} \quad (\star)$$

として算出されることが知られている[4].

(ここで、 N : 調査対象者数, $e_i = p(z_i = 1 | x_i)$: 傾向スコアとしている.)

(★) について

広告効果量の解釈

(★) の
値が正

CM接触群の方が非接触群よりも購入意向が前向きになっている割合が大きいか, 前向きになっていない割合が小さい.

(★) の
値が負

CM接触群の方が非接触群よりも購入意向が前向きになっている割合が小さいか, 前向きになっていない割合が大きい.

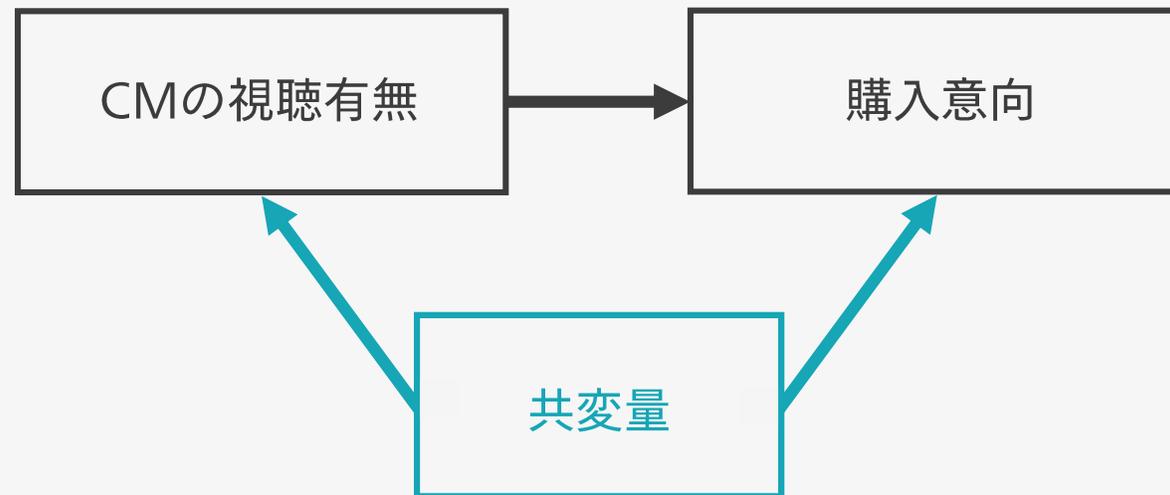


そこで, 本研究では広告効果量が

正の商品のCMを「広告効果あり」

負の商品のCMを「広告効果なし」と定義する.

本研究で用いる共変量



(★) を求めるためには上の図のような共変量を選定する必要がある. 本研究では以下を用いることとする.

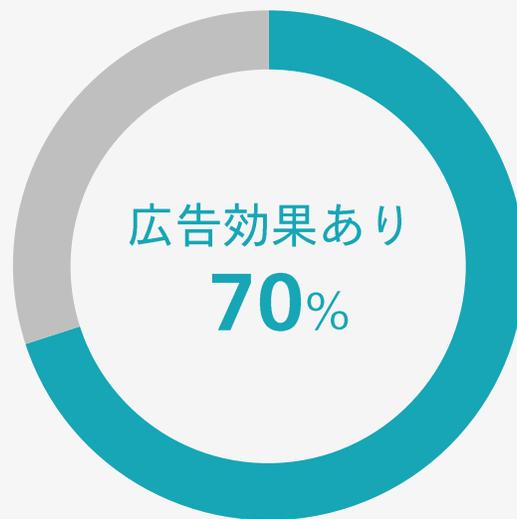
性別, 年齢, 未既婚, 家族構成, 職業,
居住形態, 世帯保有金融資産, 世帯収入

広告効果の基礎集計

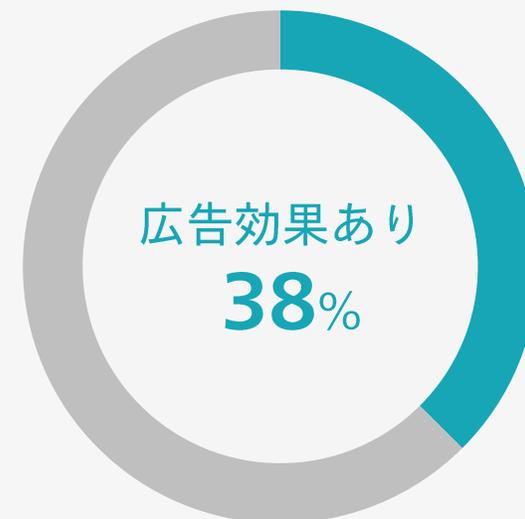
食料品



日用品



サービス類



■ 広告効果あり ■ 広告効果なし

データ加工で選別された70商品の広告効果を求めると、40商品が「**広告効果あり**」、30商品が「**広告効果なし**」となった(どの商品のCMがどのような広告効果量をとるかは補足資料を参照のこと)。

70個のCMを**食料品・日用品・サービス類**にカテゴリー分けし(内訳は食料品が26個,日用品が20個,サービス類が24個),それぞれの広告効果をまとめると上のグラフのようになった。

本分析①の流れ

ここから本分析に移る。本分析①は研究目的1を達成するための分析であり、以下の手順で行う。

Process 01

70個の商品のCMの動画データを公式HPなどから収集する。



70個のうち51個収集できた
(内訳は食料品が16個, 日用品が20個, サービス類が15個)。

Process 02

動画データからCMの質的な特徴量と量的な特徴量を定める。



次頁以降で解説する。

Process 03

CMの特徴量と広告効果の関連を多重対応分析や主成分分析を用いて視覚化する。

CMの特徴量

広告効果がある(ない)CMの特徴を抽出するために,以下のような特徴量を定めた.

変数名	質的な特徴量
MT	男性タレントは出演しているか
LT	女性タレントは出演しているか
ML	男性と女性のどちらが多く出演しているか
MN	ナレーションに男性はいるか
LN	ナレーションに女性はいるか
ST	ストーリー仕立てであるか
SI	Webページへの誘導はあるか
BV	BGMに歌詞はあるか
OV	屋外の映像はあるか
RI	企業固有の映像はあるか

変数名	量的な特徴量
using	商品を使用(飲食)する秒数
package	商品が画面に映る秒数
talent	タレントが画面に映る秒数
cutting	画面が切り替わる回数
character	画面に映る文字総数

左表のML以外の変数は,それぞれ当てはまる(=1)・当てはまらない(=0)の2値をとる. MLは,男性が多い(=1)・女性が多い(=2)・同数(=3)の3値をとる.

右表のpackageは,サービス類は無形であるため,ロゴが映る秒数をカウントした.

分析のためのデータセット

- 前頁の特徴量に加え, 以下の6値をとる変数categoryを加え, 右の2つのデータセットを作成した.
- 右上の表を用いて多重対応分析を, 右下の表を用いて主成分分析を行う.

category	内容
food_0	広告効果なしの食料品のCM
food_1	広告効果ありの食料品のCM
daily_0	広告効果なしの日用品のCM
daily_1	広告効果ありの日用品のCM
service_0	広告効果なしのサービス類のCM
service_1	広告効果ありのサービス類のCM

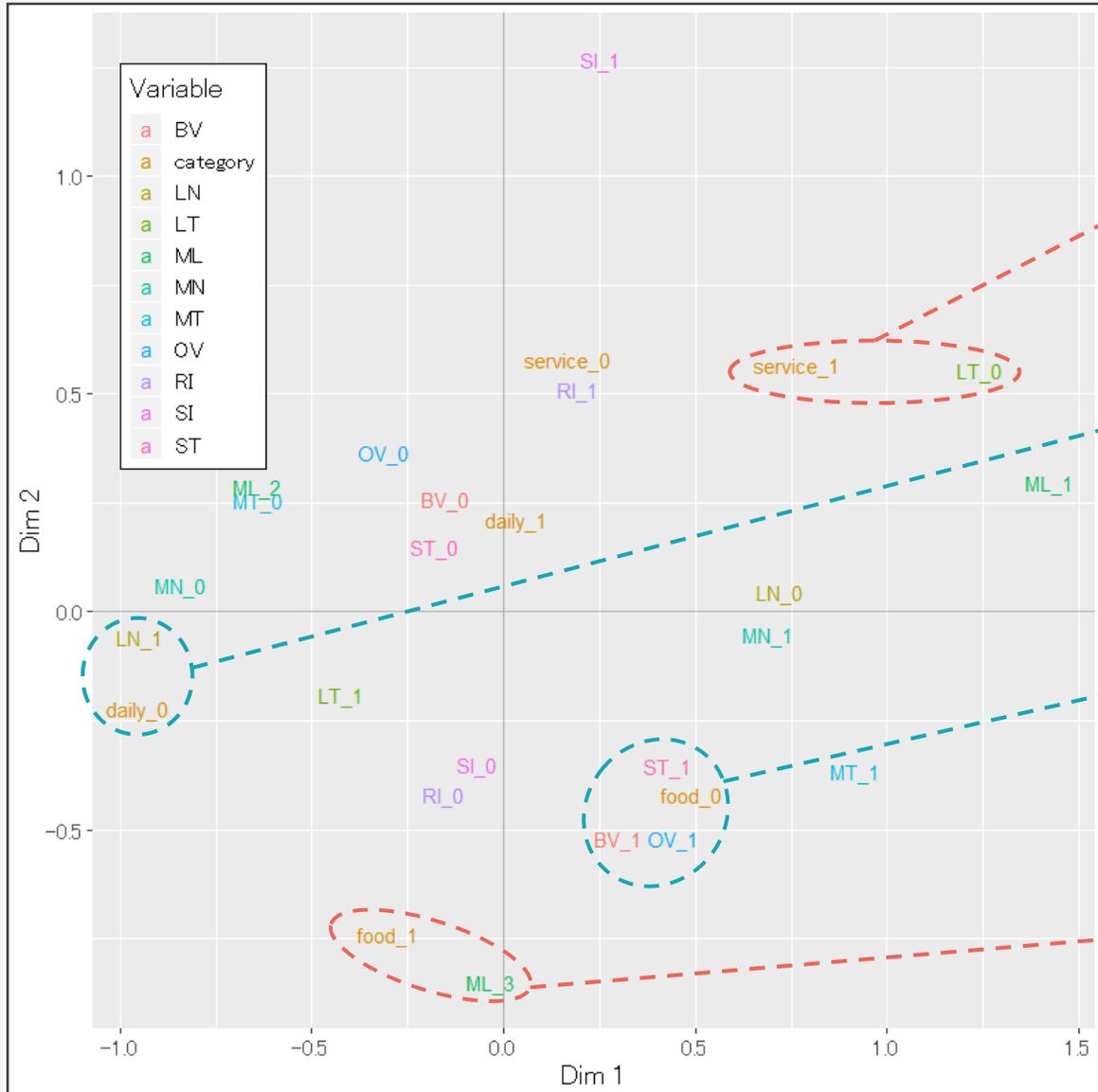


CM	category	MT	LT	...	RI
1	food_1	MT_1	LT_0	...	RI_1
2	food_1	MT_1	LT_1	...	RI_0
...
51	daily_1	MT_1	LT_0	...	RI_1



CM	category	using	...	character
1	food_1	1.17	...	299
2	food_1	2.16	...	27
...
51	daily_1	6.4	...	51

質的な特徴量を用いた多重対応分析



数量化得点の解釈

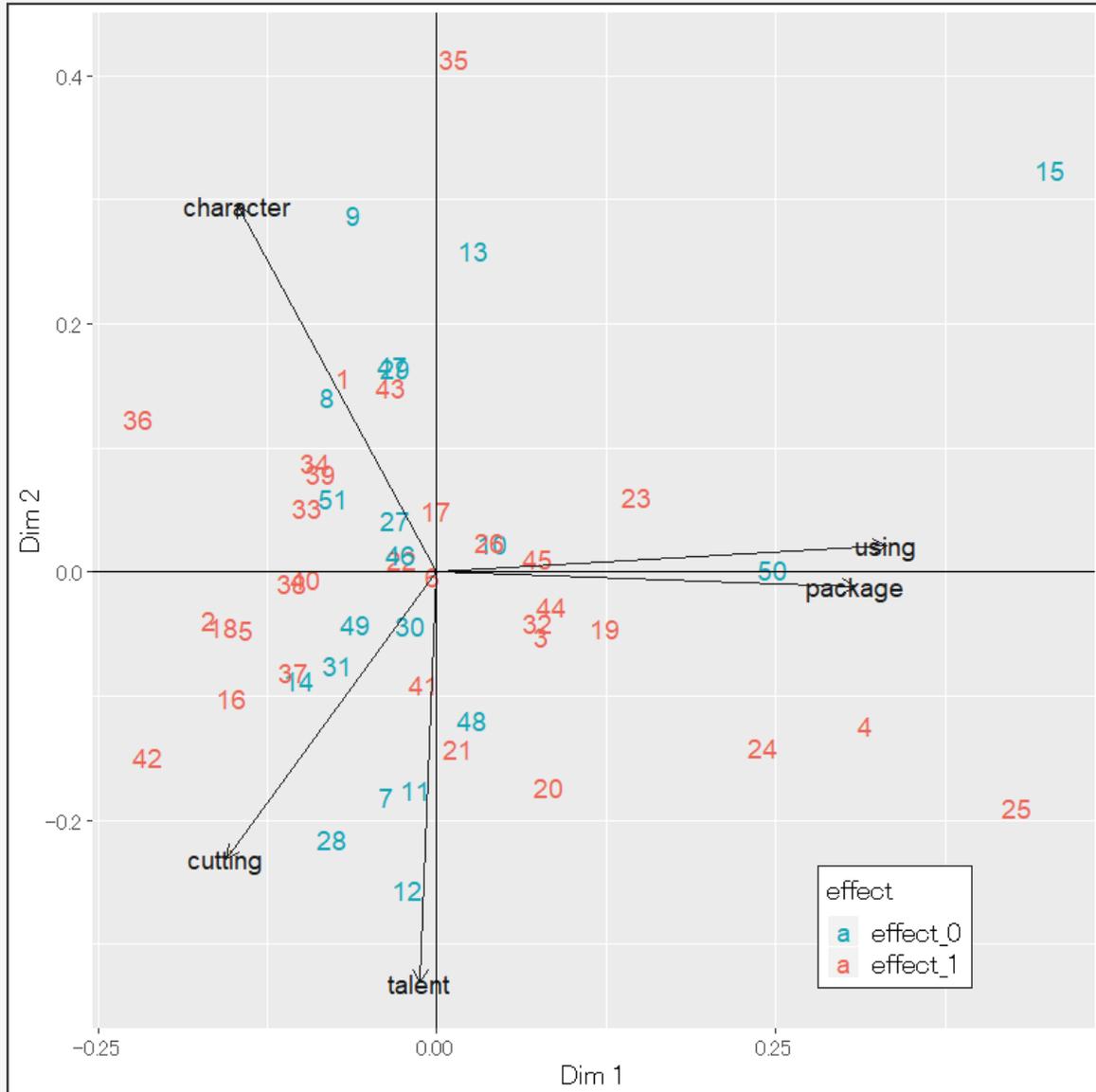
広告効果のあるサービス類のCMは、**女性タレント**が出演していない傾向が見られる。

広告効果のない日用品のCMは、**女性ナレーション**が含まれる傾向が見られる。

食料品について、**ストーリー仕立て**であったり、**BGM**に歌詞があったり、**屋外の映像**が含まれるCMは広告効果がない傾向が見られる。

広告効果のある食料品のCMは**男性と女性**で同数のタレントが出演している傾向が見られる。

量的な特徴量を用いた主成分分析



主成分の解釈

第1主成分 (Dim 1)

値が大きくなるほど、商品の描写や使用描写が増え、
値が小さくなるほど、文字やカット数が増える。



商品の視覚的情報

第2主成分 (Dim 2)

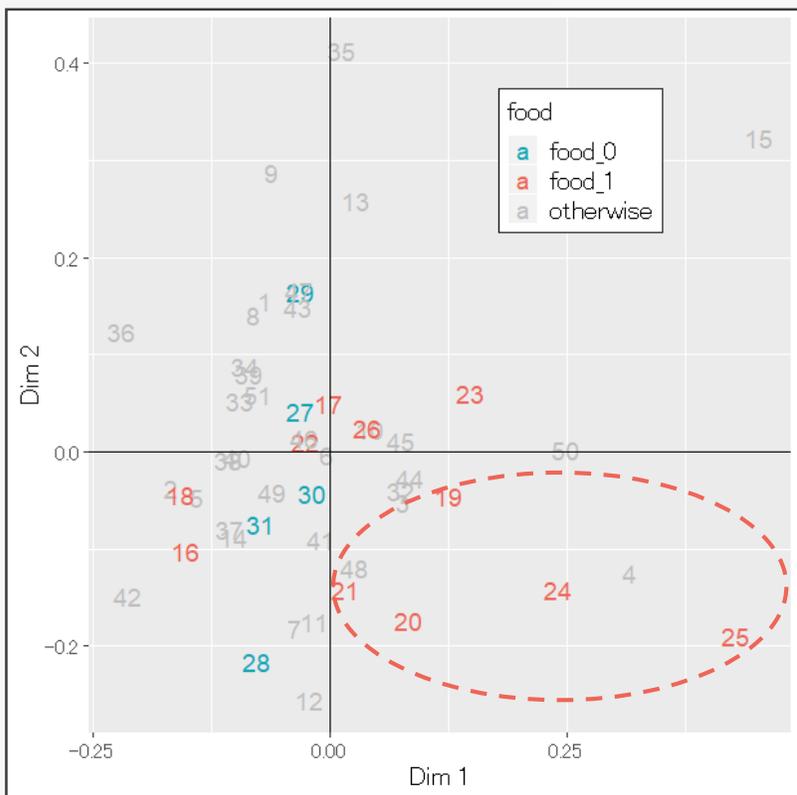
値が大きくなるほど、文字による紹介が増え、値が小
さくなるほど、タレントによる紹介が増える。



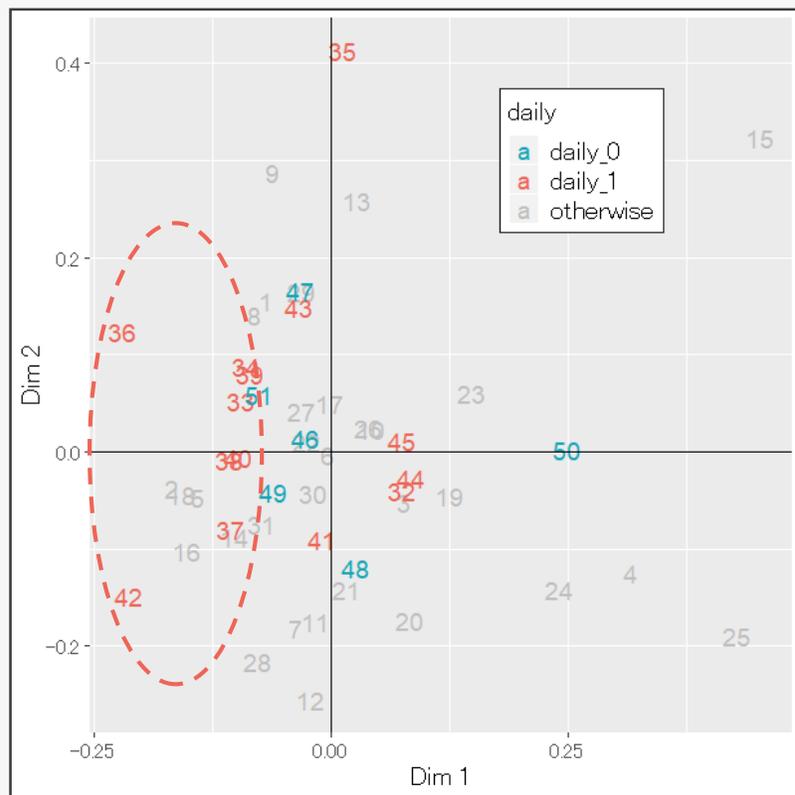
商品の紹介方法

注) 図のeffect_0は「広告効果なし」、effect_1は
「広告効果あり」を表す。

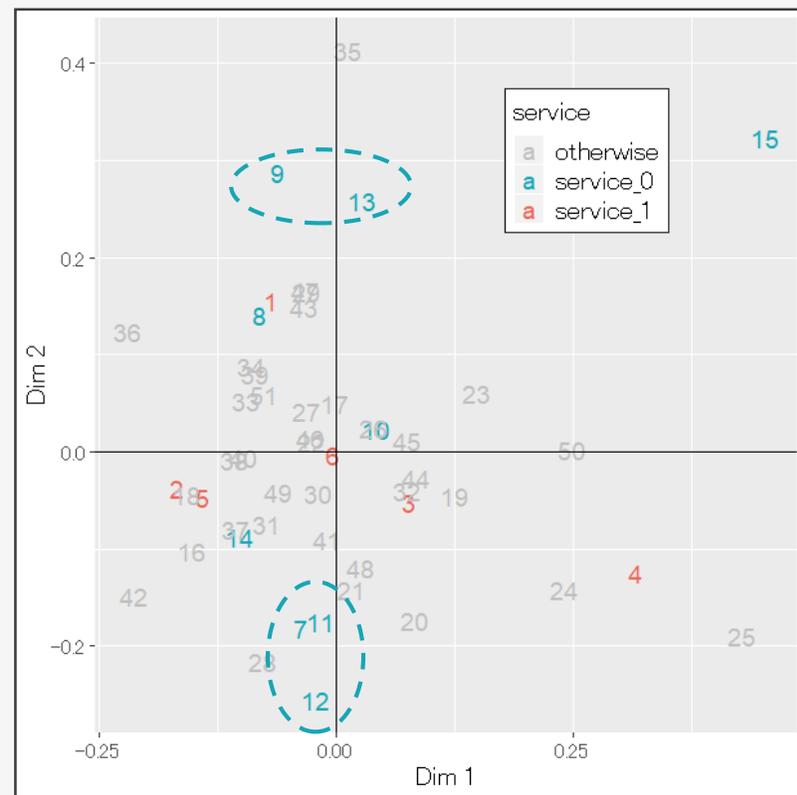
前頁の図の主成分得点を, さらにカテゴリーごとの広告効果に関して色分けすると以下ようになった。



食料品について, 商品の視覚的情報が多く, タレントが商品紹介をしているCMは広告効果がある傾向が見られる。



日用品について, 商品の視覚的情報が少なく, 文字総数と画面が切り替わる回数が多いCMは広告効果がある傾向が見られる。



サービス類について, 文字情報やタレントによる商品紹介描写が多過ぎるCMは広告効果がない傾向が見られる。

本分析②の流れ

次に本分析②を以下の手順で行う。本分析②は研究目的2を達成するための分析である。

Process 01

- CMを制作した後も、**現実的に操作可能な変数**を動画データから抽出する。
- 本研究では動画データを1秒単位で画像に区切り、**画像のRGB・HSV**を変数として用いる。

Process 02

- RGBやHSVを説明変数、広告効果量を目的変数とした**LASSO回帰**による予測モデルを作成する。
- モデル作成の際は、**交差検証法**を用いて最適なパラメータを求める。

Process 03

- モデルの正解率(※)を算出し、有用性を確認する。
 - 説明変数の解釈を行い、色彩情報と広告効果との関連を考察する。
- (※) 広告効果量の予測値について、正の値をとれば広告効果あり、負の値をとれば広告効果なしとみなして算出する。

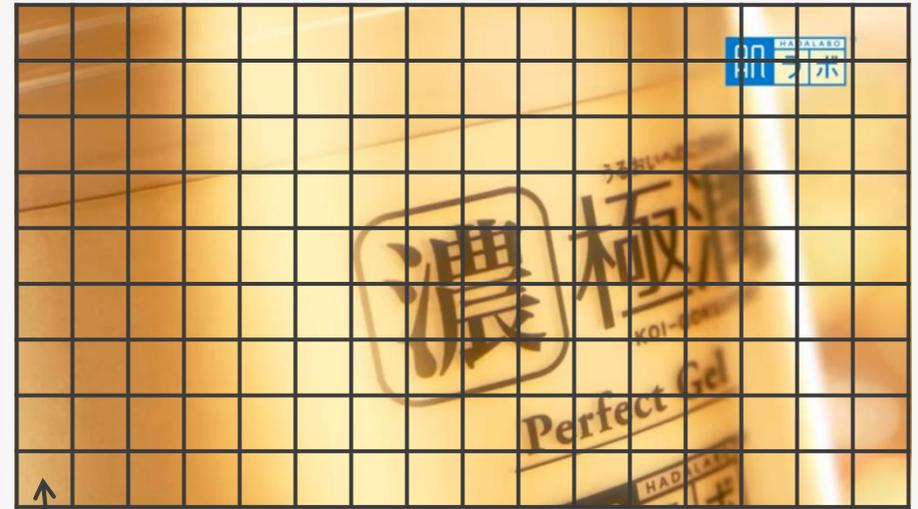
動画データから色彩情報の抽出

予測モデルの説明変数として色彩情報を選んだのは、操作性に優れていることにくわえ、CMの印象を決める重要な要因であるからである。



抽出方法

15秒のCM動画データを1秒単位で区切り、左図のような16枚の画像データ (サイズ: 1280×720) を作成する。そして画像を以下のように16×9分割し、各セルのRGBやHSVの平均を算出する。



1つのセルに80×80ピクセルが含まれる。本研究では、セルの位置情報も抽出し、例えば矢印が示すセルの座標は(9,1)を指す。

RGBとHSV

RGB

- コンピュータのディスプレイやテレビなど, 光で色を作るときのカラーモードである.
- 光の3原色である赤 (Red), 緑 (Green), 青 (Blue) の組み合わせで表現される.
- 各原色について, 度合いが弱いと0に近く, 強いと255に近い数値をとる. 度合いが弱いとより暗い色, 強いとより明るいはっきりした色になる.

赤 (R) の度合い 緑 (G) の度合い 青 (B) の度合い

0 ~ 255

0 ~ 255

0 ~ 255

(例) RGBの値が (255, 255, 255) で白, (0, 0, 0) で黒となる.

HSV

- 色を色相 (Hue), 彩度 (Saturation), 明度 (Value) の3要素で表現する方式である[5].

色相	具体的な色を定義する要素
彩度	色の鮮やかさ・濃さを表す要素
明度	色の明るさ・暗さを表す要素

色相 0  360

彩度 0  100

明度 0  100

LASSO回帰

- 取ってくる色彩情報には位置情報も付随しており、データ数に対して説明変数の数が非常に多くなる。
- そのため、通常の線形回帰モデルが使えない。そこで、本研究ではパラメータ推定と変数選択を同時に行うことができ、説明変数を解釈することができるLASSO回帰を用いる。パラメータは以下の式で推定する。

$$\vec{\hat{\beta}}_{\text{LASSO}} = \operatorname{argmin}_{\beta} \left[(\vec{y} - \mathbf{X}\vec{\beta})^T (\vec{y} - \mathbf{X}\vec{\beta}) + \lambda \sum_{j=1}^p |\beta_j| \right]$$

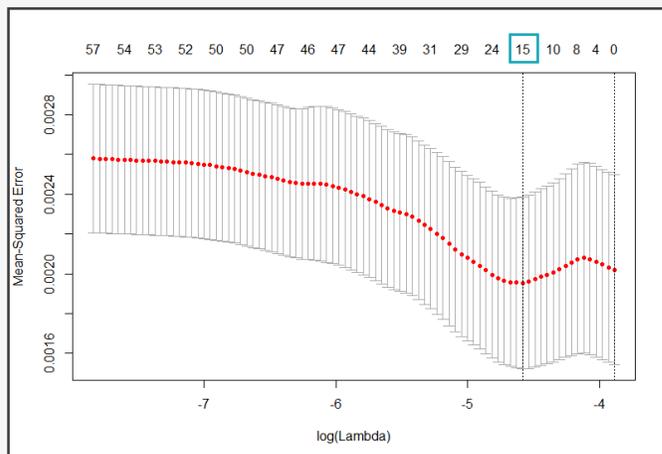
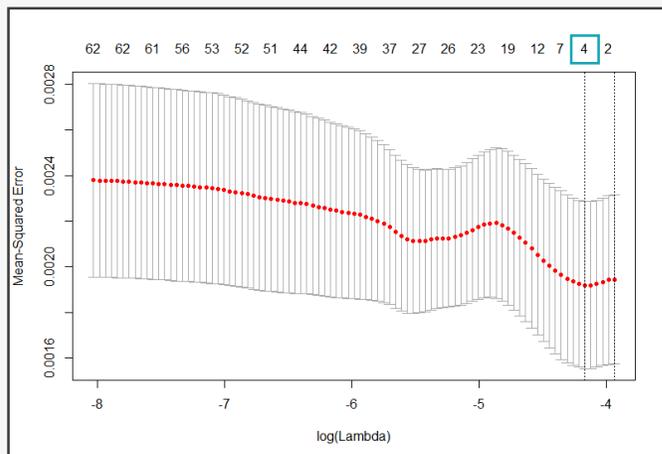
- 作成するモデルは以下の通り。ただし i, j, k は整数であり、 $0 \leq i \leq 15, 1 \leq j \leq 9, 1 \leq k \leq 16$ とする。

モデル	目的変数	説明変数
モデル1	11頁で求めた広告効果量	i 秒目の座標 (j, k) のR, G, Bの値
モデル2	同上	i 秒目の座標 (j, k) のH, S, Vの値

(※) 説明変数は両モデルとも画像データ16枚分の位置情報16×9個の3変数なので、計6912 (=16×16×9×3) 個である。

推定手順と結果

まず交差検証法を用いて、推定値と観測値の平均二乗誤差が最小となるように λ を決定する。



↙
 λ の値と平均二乗誤差の関係を示した図. 上がモデル1, 下がモデル2のものである.
 ↙

左図より、モデル1では4、モデル2では15の変数が選択された. 最適な λ を用いてモデルを作成した結果とモデルの正解率は以下の通り.

モデル1の結果

変数名	β
(Intercept)	0.00459
0秒目:(3,9)のR値	-3.13E-05
10秒目:(9,1)のR値	3.09E-05
11秒目:(9,3)のG値	3.78E-05
12秒目:(9,3)のG値	1.46E-05

モデル2の結果

変数名	β
(Intercept)	-0.03379
0秒目:(3,8)のH値	2.29E-05
0秒目:(4,8)のH値	1.06E-06
0秒目:(4,11)のH値	2.68E-05
1秒目:(4,2)のH値	4.82E-05
1秒目:(4,3)のH値	3.56E-06

5秒目:(9,10)のH値	3.69E-06
13秒目:(3,13)のH値	8.84E-06
2秒目:(5,7)のS値	3.86E-05
2秒目:(5,16)のS値	0.00013
4秒目:(2,3)のS値	0.00015
11秒目:(3,12)のS値	2.25E-05
0秒目:(3,9)のV値	-3.23E-05
10秒目:(9,1)のV値	0.00012
11秒目:(9,3)のV値	0.00019
12秒目:(9,3)のV値	0.00014

正解率

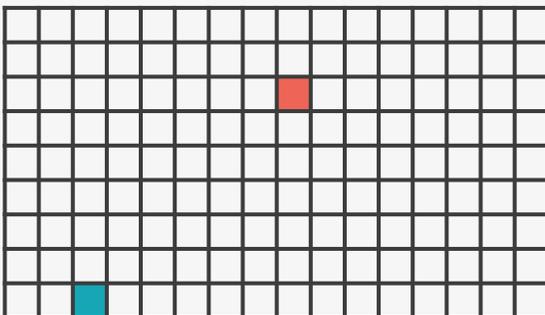
モデル1	0.63
モデル2	0.76

結果の考察

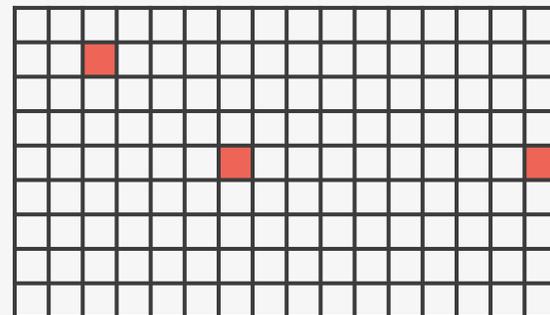
モデルの正解率について、データ数の不足や説明変数の限定などの理由からNRIの予測モデルに比べると劣るが、色彩情報だけを用いて作成したことを考慮すると及第点のモデルが作成できたと考える。

<h3>モデル1の解釈</h3>	<p>CMの序盤に(i)の赤い部分の赤 (R) の度合いが高いとよくない。これは、赤色が注意や危険を表す色であるため、最初に目に入りやすい部分にこの色を配置することで悪い印象を与えるからだと考えられる。一方で、CMの終盤に(i)の青い部分の緑 (G) の度合いが高いとよい。この部分は、CMの終盤に企業のロゴや文字による補足情報が含まれることがあり、それらは安全や安定といったイメージを与える緑との相性がよいことから変数として残ったものと考えられる。</p>
<h3>モデル2の解釈</h3>	<p>CMの初頭に(ii)の赤い部分の彩度 (S) を大きくするとよい。序盤は初頭効果が働くため、見る人の印象に残りやすい。そこでタレントが映りやすい位置 ((2,3), (5,7)) や、目線の高さで左から右への視線の移動の先にある位置 ((5,16)) の色を鮮やかにすることによって好印象をもたらすCMになるのではないかと考えられる。また、CMの終盤に(iii)の青い部分の明度 (V) を大きくするとよい。この部分はCMの終盤に視線が動く位置であり、最後に暗い部分を見ることでそのCMの印象に悪影響を与えるものと考えられる。</p>

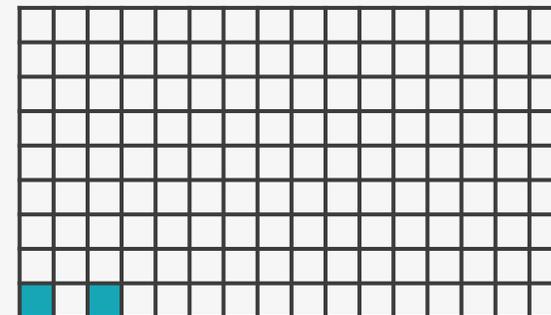
(i)



(ii)



(iii)



CM制作の指標提言

本研究のまとめとして、まず商品カテゴリーごとにCM制作の指標を提言する。



食料品のCMの広告効果について、男女同数のタレントを出演させて、当該商品の描写とタレント (が飲食を行う) 描写の秒数をできる限り長くするとよい。また、ストーリー仕立てであったり、BGMに歌詞が付いていたり、屋外の映像が含まれているとよくない。



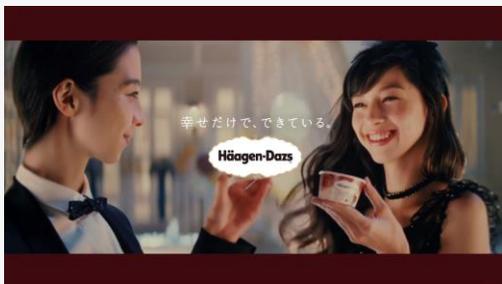
日用品のCMの広告効果について、当該商品の描写や使用描写の秒数を短くして、画面に映る文字総数と画面が切り替わる回数を多くするとよい。また、ナレーションに女性の声が含まれるとよくない。



サービス類のCMの広告効果について、女性タレントを出演させないとよい。また、文字情報もしくはタレントの描写が過多であるとよくない。さらには、企業固有の映像がCMに含まれるのもよくない。

本研究で用いたCMの具体例

広告効果がある食料品のCM



「ハーゲンダッツ」のCM. 出演タレントの男女2人が常に商品を手にしており、パッケージの描写が長い(ただし男性タレントは女性タレントの男装).

広告効果がない食料品のCM



「ザ・プレミアム・モルツ」のCM. 出演タレントが人通りの多い夜の雑踏を歩く描写が長く、パッケージや飲酒の描写が短い。また、BGMに歌詞が含まれる。

広告効果がある日用品のCM



「ディアボーテ HIMAWARI」のCM. パッケージの描写よりも出演タレントが向日葵であふれる花屋で会話する描写が長く、画面が切り替わる回数も多い。

広告効果がない日用品のCM



「クラリチンEX」のCM. スポーツカーが荒野を駆け抜ける描写が長く、出演している女性タレントがナレーションも兼ねている。

広告効果があるサービス類のCM



「タウンワーク」のCM. 出演タレントの衣類に商品のロゴマークがついているため、ロゴが画面に映る時間が長く、女性タレントが出演していない。

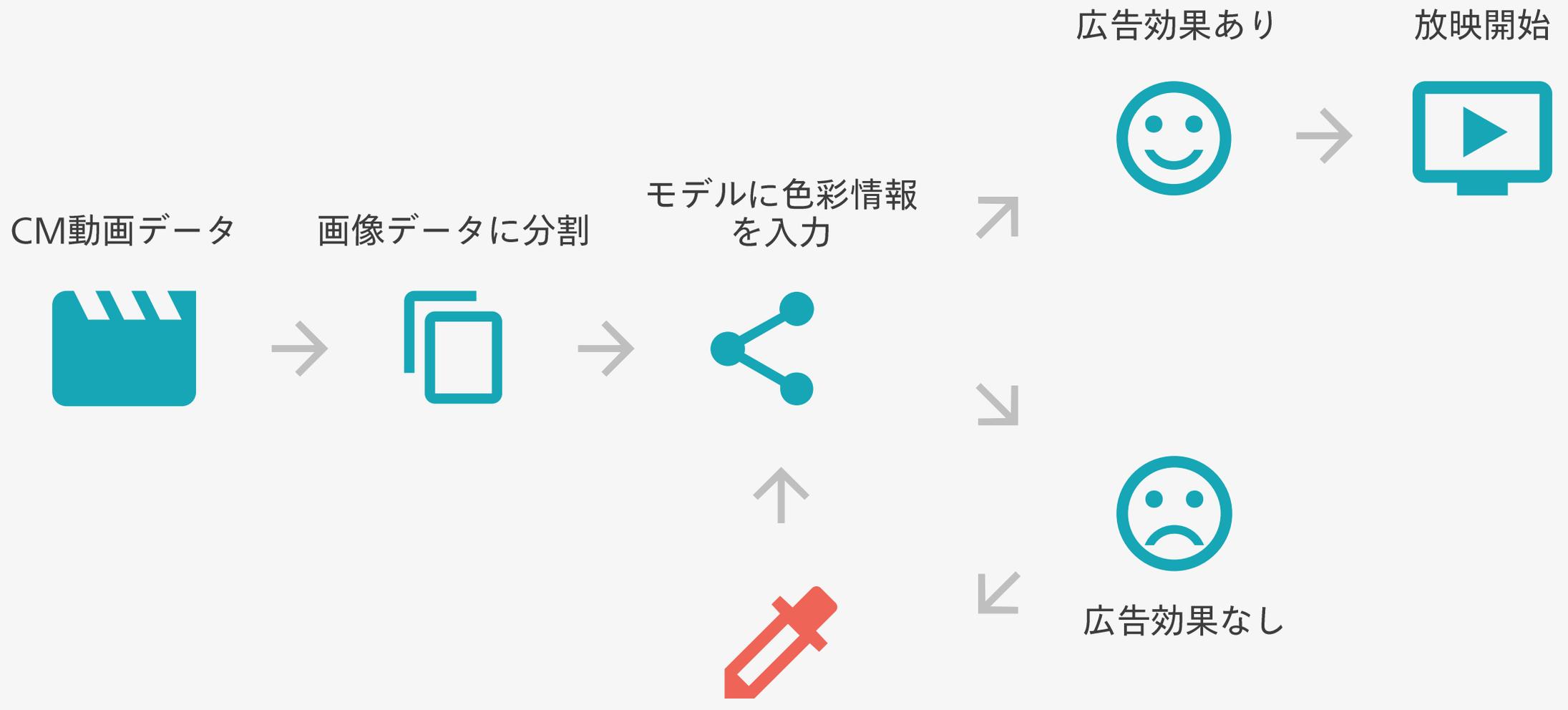
広告効果がないサービス類のCM



「SMBC日興証券」のCM. SMBCデビットの説明のための文字情報が多く、上図のような企業固有の映像が含まれている。

予測モデルの活用

最後に, 作成したモデルを実際にどのように活用するか図解する.



24頁の結果などを参考に
色彩情報を操作

補足資料1

各商品の広告効果量

「広告効果あり」のCMは赤、「広告効果なし」のCMは青で色分けしている。

商品は食料品, 日用品, サービス類の順に並んでおり, 広告効果量の列の色が青から赤に変わる行がカテゴリーの境目である。

商品名に◎が付いているものは, 動画データまで収集できた商品である。

商品名	広告効果量
レッドブル (Red Bull)	0.0829
からだすこやか茶W◎	0.0587
ケンタッキーのフライドチキン	0.0566
綾鷹◎	0.0566
リポビタミンD◎	0.0529
伊右衛門◎	0.0470
明治プロビオヨーグルトR-1 ドリンクタイプ◎	0.0462
ヤクルト400	0.0402
オロナミンC◎	0.0261
スーパードライ	0.0212
ゴールデンカレー◎	0.0170
すき家	0.0140
ハーゲンダッツ ミニカップバニラ◎	0.0124
マクドナルド	0.0109

商品名	広告効果量
お~いお茶◎	0.0084
Newヤクルト	0.0052
ガスト◎	0.0001
黒烏龍茶◎	-0.0001
バーモントカレー	-0.0039
丸亀製麺	-0.0100
一番搾り◎	-0.0113
生茶	-0.0199
キットカット◎	-0.0200
十六茶	-0.0512
ザ・プレミアム・モルツ◎	-0.0579
きのこの山 たけのこの里◎	-0.0824
いち髪 スタイリングシリーズ◎	0.0903
Essential (エッセンシャル) の シャンプー・コンディショナー◎	0.0795

補足資料2

商品名	広告効果量
ジュレーム◎	0.0724
ドクターシーラボ アクアコラーゲンゲル◎	0.0616
ディアボーテ HIMAWARI◎	0.0513
モイストダイアンボディソープ◎	0.0429
ストナリニS◎	0.0428
ダヴボディソープ◎	0.0337
アレグラFX◎	0.0273
パブロン鼻炎カプセルSα◎	0.0162
肌ラボ 極潤 パーフェクトゲル◎	0.0127
ニベアボディソープ◎	0.0120
dプログラム アレルバリアシリーズ◎	0.0030
プリマヴィスタ きれいな素肌質感 パウダーファンデーション◎	0.0165
クラリチンEX◎	-0.0029
ラメランス◎	-0.0083

商品名	広告効果量
ソフティモ ナチュサボン◎	-0.0125
マキアージュ ドラマティックパウダリーUV◎	-0.0205
専科 ボディソープ◎	-0.0328
ラックスルミニーク◎	-0.0390
三菱東京UFJ銀行◎	0.0977
Renta!◎	0.0574
ホットペッパービューティー◎	0.0467
JALの先得割引	0.0416
タウンワーク◎	0.0403
ローソン	0.0121
ミュゼプラチナム◎	0.0090
三井住友銀行◎	0.0042
三菱UFJ信託銀行◎	0.0036
マイナビ◎	-0.0043

商品名	広告効果量
an	-0.0053
SMBC日興証券◎	-0.0054
大和証券◎	-0.0081
東急リバブル	-0.0094
アットホーム◎	-0.0199
いい部屋ネット◎	-0.0201
三井住友信託銀行	-0.0296
バイトル◎	-0.0394
ANAの旅割	-0.0516
ゼクシィ	-0.0519
コミックシーモア◎	-0.0587
野村証券◎	-0.0597
セブンイレブン	-0.0849
LIFULL HOME'S◎	-0.1088