

YouTube 利用者の視聴数モデルの構築

1X12C086-1 行川大輝
指導教員 大野高裕

1. はじめに

2005年2月に設立された YouTube は2016年国内において3580万人が活用しており、影響力のあるメディアとして台頭している。多くのアーティストが楽曲を YouTube 上で無料公開しており、知名度を上げるためのマーケティング活動の場所としても大変重要な場所になっている。

本研究では、YouTube で楽曲の視聴状況が時系列でどのように変化するか把握するモデルを構築することで、各ユーザーの状態推移を明らかにすることを目的とする。

2. 従来研究[1]

顧客の状態が時間経過によって対象の状態が変化するものをモデル化する方法として微分方程式がある。そして微分方程式を用いて、状態の推移を表現した代表的なものとして中桐裕ら[1]の社会的ブームの微分方程式モデルがある。

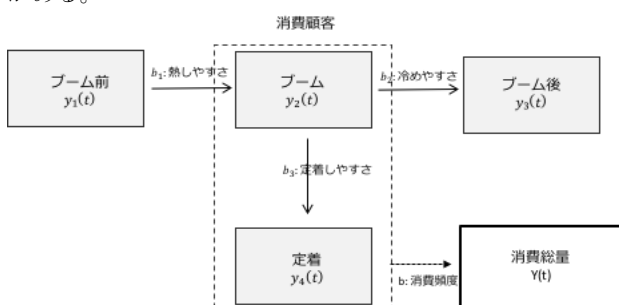


図 1. 従来モデル[1]

中桐ら[1]はこのモデルにおいて社会的なブームを明らかにすることを目的としており、日用品の購買行動に注目してモデルを構築している。日用品は再購買されやすい商品のため、一度定着したら商品を購入し続けることが多いと考えられるが、今回のテーマである YouTube のように顧客が定着しづらい商品においては、定着後に視聴をやめることを考慮に入れたモデルを構築する必要がある。

3. 提案モデル

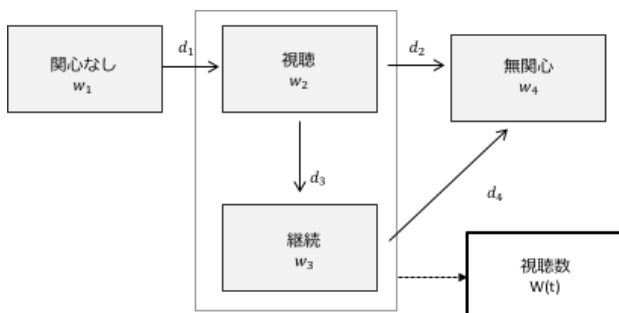


図 2. 提案モデル

本研究では図2で示されるモデルを提案する。このモデルでは YouTube ユーザーの推移を表現しており、視聴を始めた人が飽きるまでという全体の流れを表現している。

それぞれの状態は、

- w_1 : 《関心なし》: 対象への関心がない状態
- w_2 : 《視聴》: YouTube で対象アーティストを認知
- w_3 : 《継続》: 対象への関心をもち活用している状態
- w_4 : 《無関心》: 対象への関心がなくなった状態と定義する。

また、 w_1, w_2, w_3, w_4 は人数を表し、

$$w_1(t) + w_2(t) + w_3(t) + w_4(t) = S \text{ が成り立つ。}$$

S は全顧客数を表している。

状態遷移のパラメーターは、

- d_1 : 認知し始める割合
- d_2 : 関心を無くす割合
- d_3 : 継続して視聴する割合
- d_4 : 関心をなくし活用をやめる割合と表現する。

図より微分式を導出し、その式に初期条件を与えると、

$$w_1(t) = (1-k)e^{-d_1(t-T)}$$

$$w_2(t) = \frac{d_1(1-k)}{D_1}(-e^{-d_1(t-T)} + e^{-(d_2+d_3)(t-T)})$$

$$w_3(t) = \frac{d_1 d_3(1-k)}{D_1 D_2} e^{-d_1(t-T)} + e^{-(d_2-d_3)(t-T)} + \left(\frac{d_1 d_2}{D_2(d_1-d_4)}(1-k)(e^{-d_1(t-T)} - e^{-d_4(t-T)}) + k e^{-d_4(t-T)}\right)$$

$$w_4(t) = \left(d_3 + \frac{d_2 d_4}{D_2}\right) \left(\frac{1-k}{D_1(d_2+d_3)}\right) E + \left(\frac{d_2(1-k)}{D_2(d_1-d_4)}\right) (-d_4 e^{-d_1(t-T)} + d_1 e^{-d_4(t-T)}) - k e^{-d_4(t-T)} + \left(\frac{d_2(1-k)}{D_2}\right) \left(1 + \frac{d_4}{d_2+d_3}\right) + \frac{1}{d_2+d_3} \{d_3(1-k)\} + k$$

$$D_1 = (d_1 - d_2 - d_3)$$

$$D_2 = (d_4 - d_3 - d_2)$$

$$E = (d_2 + d_3)e^{-d_1(t-T)} - d_1 e^{-(d_2-d_3)(t-T)}$$

が導き出される。

この式と実数値を元にパラメーター d_1, d_2, d_3, d_4 を推定し、パラメーターは Excel のソルバーを用いて推定する。

4. データと分析手法

4.1 使用データ

Google 社が提供している Artist Insight[2]を活用する。モデルの分析には2016年上半期のレコチョクランキング[3]の30位以内からランダムに選出した bump of chicken(2016/8/26~2016/10/28) perfume(2015/10/9~2015/12/25) superfly(2015/11/13~2016/1/22) radwinps(2016/8/1~2016/11/30) のデータを用いる。

モデルの当てはまりの指標は、実データを目的変数、推定値を説明変数とする回帰式によって導出される決定係数 R^2 を用いる。

4.2 分析結果

図 3,4 に bump of chicken を例として提案モデルと従来モデルのグラフを提示する。

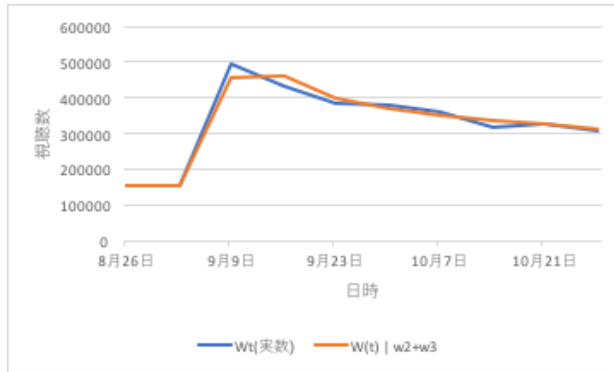


図 3.提案モデル($R^2=0.917$)

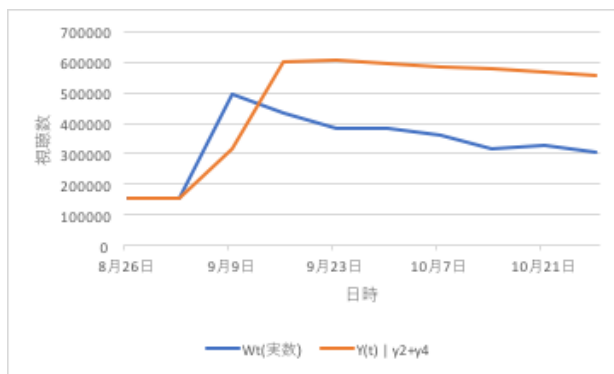


図 4.従来モデル($R^2=0.228$)

提案モデルとの方が目測でのグラフの合致度が良く、かつ決定係数の値も優れた結果が導出された。

5.考察

表 1 の決定係数に注目すると、どれも高い数値を示しており、モデルの有用性を示している。次に各パラメーターに注目する。 d_4 はほぼ 1 であり、1 度興味を持って動画を視聴しても一定の時間が経つと関心を失い沈静に状態が移ることがわかる。

理由として、

- 1.時間とともにアーティストの曲に飽きている。
- 2.有料聴き放題サービスなどの別サービスの活用などが考えられる。

表 1.パラメーターまとめ

	bump of chicken	perfume	superfly
d_1	0.00516	0.00640	0.000738
d_2	0.865	0.779	0.741
d_3	0.716	0.484	0.651
d_4	0.990	1.00	1.00
R^2	0.968	0.875	0.956

次にモデルが当てはまらない例に注目する。

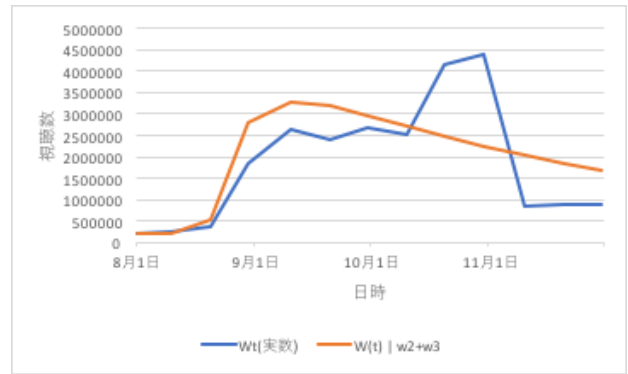


図 5.提案モデル($R^2=0.130$)

radwips(2016/8/1~11/30)で検証した場合、提案モデルの当てはまりが悪かった。モデルが当てはまらなかった場合と比較してみると、一度視聴数が大きく伸びた後に、もう一度視聴数が伸びている点が大きく違う。モデルでうまく表現出来ない理由としては、モデルで表現しているパラメーターが定数であることが考えられる。何らかの要因で d_1 の値が急激に上昇、もしくは d_2, d_3 の値が減少したとしても、短い期間ではその動きをモデルとして捉えることが出来ず曲線カーブを描く推移しか表現することが出来ない。よって今後の課題としてはパラメーターを変数として扱う方法を考える必要がある。

6.おわりに

本研究では微分方程式モデルを用いて

w_1 :《関心なし》:対象への関心がない状態

w_2 :《関心あり》:対象への関心を持った状態

w_3 :《継続》:対象への関心をもち活用している状態

w_4 :《沈静》:対象への関心がなくなった状態

という 4 つの状態が顧客が存在することが明らかになった。

また各状態ごとに、

d_1 :関心を持ち始める割合

d_2 :関心を無くす割合

d_3 :実際に活用する割合

d_4 :関心をなくし活用をやめる割合

と表現出来ることが明らかになった。

しかし、パラメーターの値が急激に変化する時は、パラメーターを定数ではなく、変数として扱うことができるモデルを構築することが今後の課題として上げられる。

参考文献

[1] 中桐裕子, 栗田治:「社会的なブームの微分方程式モデル」,日本オペレーションズ・リサーチ学会和文論文誌 vol47, pp. 83-105,(2004)

[2]artistinsight

<https://www.youtube.com/yt/artists/ja/insights-artist.html>(最終閲覧日 1/8)

[3]レコチョク アーティストランキング

<http://recochoku.jp/special/100451/>

(最終閲覧日 1/14)