

平成 29 年 6 月 2 日

各位

会社名 株式会社ソケット  
代表者名 代表取締役社長 浦部 浩司  
(証券コード：3634)

---

## ソケット、ドラマのあらすじから感情をスコア化、 2017 年春ドラマの視聴率を予測

---

株式会社ソケット（本社：東京都渋谷区、代表取締役社長：浦部浩司、以下「ソケット」）は、自社のMSDB（※）より抽出した、2003 年以降のプライムタイム放送の連続ドラマデータから、感情分析エンジンを活用して、あらすじから感情スコア値を算出、視聴率との関係性を分析し、さらにドラマが表現する“感情”をベースとした視聴率予測モデルを構築し、実際に現在放送中の今期春ドラマの視聴率予測を行ったレポートを公開いたしました。

ソケットでは、音楽、映像、書籍をはじめとしたエンターテインメントに関するメタ情報と、感情分析エンジンによる人の“感情”をキーとした分析予測から、商品開発・調達、マーケティング支援に向けた実用化を目指しております。

（※）MSDB（メディアサービス・データベース）とは、ソケットが開発した音楽、映像、書籍、放送、人物、施設、一般商品情報などを体系的かつ特徴情報を詳細に分類したデータベースであり、人の感性や感情を捕捉した「感性メタデータ」をキーとしたプロファイリング、アナリティクス、パーソナライズ、レコメンド、マーケティングサービスを実現します。

### 2017 年春ドラマ視聴率を予測する シリーズ第 1 回

前回までは歌詞データを元に、ソケット感情分析に焦点を当ててご紹介してきましたが、今回は、あらゆる年代、あらゆる人の想像力を掻き立てる、身近なエンターテインメント、各局放映中の“春ドラマの視聴率予測”をお届けしたいと思います。

視聴率と言えば、2000 年代に入り、若者の「テレビ離れ」が叫ばれるようになり、さらに近年では、録画視聴や見逃し配信などによる視聴の多様化により、「視聴率」という指標自体がたびたび議論されるようになりました。視聴率の定義、捉え方は時を経て変化してきているかもしれませんが、それでも、テレビ番組での紹介による爆発的ヒット現象や、Twitter のトレンドを見ても、メディアとしてのテレビの影響度自体、揺らいでいないと感じることも多いのではないのでしょうか。

「大ヒットは予期せぬところから…」とよく言われますが、あえて、その難解な領域に、データを活用した予測科学で踏み込んでみたいと思います。

---

## 連続ドラマをデータから考察

---

まず、対象データを、2003年以降のプライムタイム（19時～22時台）の全国ネット主要4局の連続ドラマデータとして、MSDBより抽出することとします。

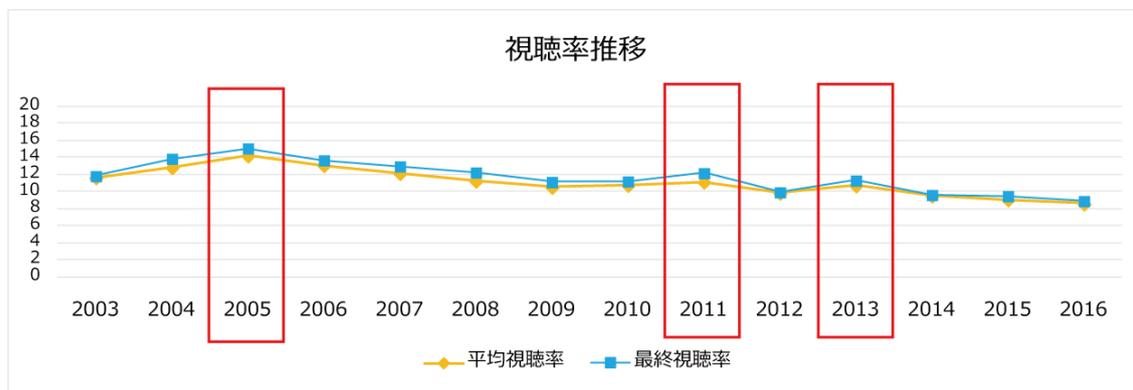
■2003年1月期～2017年4月期までの地上波プライムタイム（19時～22時台）全国ネットの民放4局（日本テレビ、TBS、フジテレビ、テレビ朝日）の連続ドラマ

■上記に付随する、放送曜日・時間帯、ドラマジャンル、各話視聴率、平均視聴率、主演、出演者、プロデューサー・監督・脚本などの付帯情報全般、あらすじデータ

■MSDBより抽出した上記データのうち、ドラマあらすじから感情分析エンジンで感情スコア値を算出

本題の視聴率予測に入る前に、簡単に対象時期の視聴率推移をご紹介しておきたいと思います。

こちらが2003年から昨年2016年までの、約14年間の視聴率推移です。



2005年をピークに、想像通り、右肩下がりとなっています。

2005年は、土9「ごくせん」（主演：仲間由紀恵）、月9「エンジン」（主演：木村拓哉）、木10「電車男」（主演：伊藤淳史）、金10「花より男子」（主演：井上真央、松本潤）をはじめ、10作品が最終回視聴率20%越えという、ヒット作に恵まれた年でした。また、ギリ貧の中、

少しだけ持ち直した 2011 年は、水 10「家政婦のミタ」、日 9「JIN-仁-（完結編）」が、2013 年には日 9「半沢直樹」が、それぞれ大ヒットし、この年の視聴率を牽引しています。

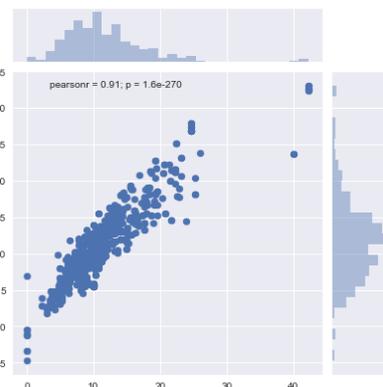
2003～2016年 連続ドラマ最終回視聴率上位10									
放送年	季節	放送局	放送曜日	放送時間	ドラマタイトル	平均視聴率	最終回視聴率	ジャンル	主演
2013	夏	TBS	日	21:00～	半沢直樹	28.7%	42.2%	経済	堺雅人
2011	秋	日テレ	水	22:00～	家政婦のミタ	25.2%	40.0%	ホーム	松嶋菜々子
2003	冬	TBS	日	21:00～	GOOD LUCK!!!	30.4%	37.6%	ビジネス	木村拓哉
2005	冬	日テレ	土	21:00～	ごくせん	28.0%	32.5%	学園	仲間由紀恵
2004	冬	フジ	木	22:00～	白い巨塔 第二部	26.2%	32.1%	医療	唐沢寿明 江口洋介
2007	冬	TBS	日	21:00～	華麗なる一族	24.4%	30.4%	ビジネス	木村拓哉
2004	冬	フジ	月	21:00～	プライド	25.2%	28.8%	恋愛	木村拓哉 竹内結子
2007	冬	TBS	金	22:00～	花より男子2（リターンズ）	21.6%	27.6%	恋愛	井上真央 松本潤
2014	秋	テレ朝	木	21:00～	ドクターX～外科医・大門未知子～	22.9%	27.4%	医療	米倉涼子
2008	春	フジ	月	21:00～	CHANGE	22.1%	27.4%	政治	木村拓哉 深津絵里

## 連続ドラマ視聴率予測モデルを構築する

それでは早速、最終回視聴率の予測モデルを構築していきたいと思います。

まずは一般的な重回帰分析で初回～5 週目までの視聴率データを元に予測してみると、誤差±4 ポイント前後で 7 割程度は実際の最終回視聴率の予測ができました。

ただ、肝心のヒットドラマ、または大低迷ドラマに関しては予測値が大きく外れる結果となりました。こちらを踏まえて、ドラマのあらすじから算出したソケットズ独自の感情スコア値を含め、ソケットズにしかできない予測モデルを構築していくことにします。



参考① 重回帰分析結果

(縦軸：予測結果 横軸：最終話実測値)

最終回視聴率を予測するに当たり、影響度の高いデータを検証していきます。

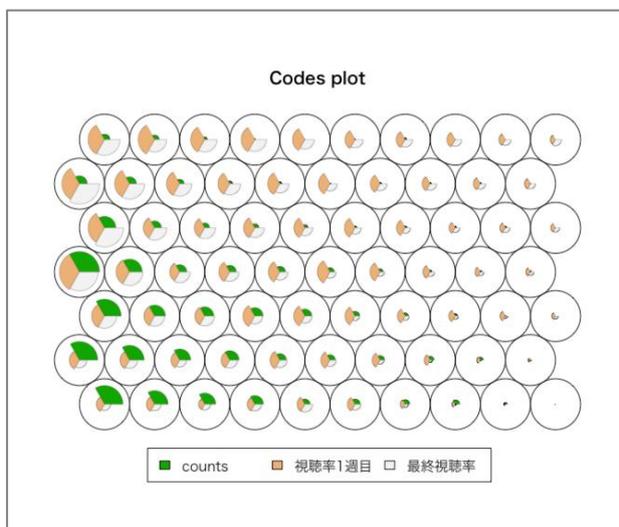
まずは、そろそろお馴染みになりつつあるソケットズ感情スコアについてですが、今まで同様、10 種類の感情でそれぞれスコア値を算出しました。歌詞の時とはキーワードの受け取り方が少し変わりますので、ワーディングはドラマのあらすじ仕様にしてあります。

幸福・愛情	嬉しい・楽しい	安らぎ・信頼	哀しみ	昂ぶり・興味
希望	不安・恐れ	怒り・苛立ち	嫌い・不愉快	退屈・うんざり

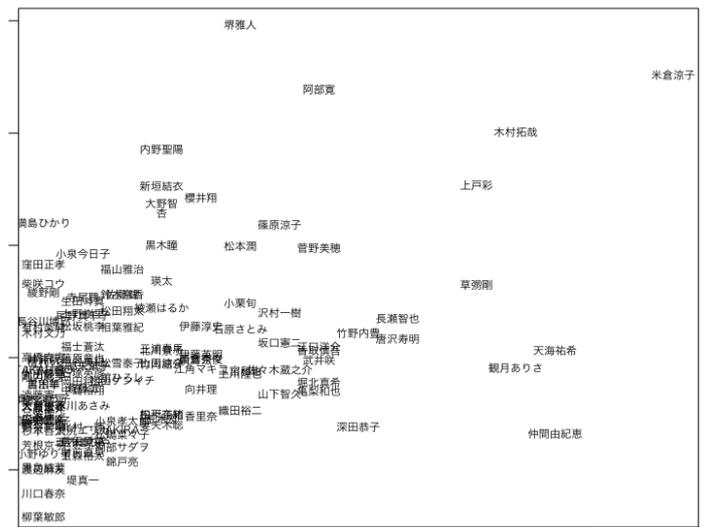
この感情スコアでいえば視聴率と最も相関が高かったのは「昂ぶり・興味」でした。逆に「退屈・うんざり」が1番影響度が低いスコアとなりました。

ちなみに、今回MSDBから抽出して利用したドラマのあらすじデータは、話数ごとではなく、各放送局のオフィシャルな番組紹介内容を元としたドラマ紹介的なあらすじ内容となっています。今回はこちらのデータから感情スコア値を算出していますが、このあらすじデータのクオリティにより、感情スコア値の精度が左右されるといっても過言ではありませんので、今後、予測モデルのチューニングとあわせて、あらすじデータの精度UPも並行して調整していきたいと思えます。

そのほか、MSDBから抽出したメタ情報、付帯情報などから、傾向値を俯瞰してみました。



参考② 主役回数と視聴率の傾向をプロット



参考③ 出演回数と視聴率の相関

(縦軸：出演回数 横軸：視聴率)

さらに近年、放送中の番組内容がツイートされることが多い、TwitterでのTweet数も視聴率との相関関係が深いものと仮定し、予測モデルに取り込むことにしました。

その結果、以下のプロトタイプ版が出来上がりました。

【視聴率予測モデル プロトタイプ】

$$x = \frac{P(m)+P(m+1)+P(m+2)}{3} \dots\dots\dots (1)$$

今回は、あるドラマの3週目までの視聴率をPとし、単純移動平均をxとする。  
mは1話目、m+1は2話目、m+2は3話目である。

あるドラマのTwitterの情報に関しては以下のように活用する。

$$T = \frac{Tw(m)+Tw(m+1)+Tw(m+2)}{3} \dots\dots\dots (2)$$

Tは1週目から3週目のあるドラマの1話ごとの放映週前後のドラマタイトル・愛称を含む Tweet 数の単純移動平均である。

以上の(1)(2)を元にプロトタイプのモデルを記述する。  
Yを求めることで各ドラマの視聴率を算出する。

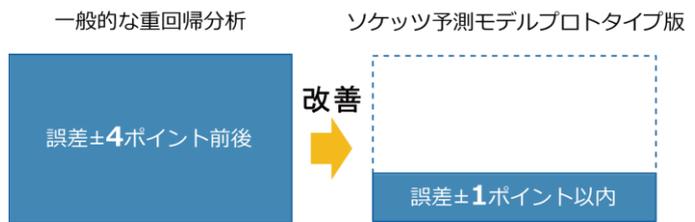
$$Y_i = x_i * \left( 1.1 - (\log(T_i)) * \frac{\sum_{i=1}^n \text{感情スコア} P_i \text{感情スコア} N_i}{x_i} \right) \dots\dots\dots (3)$$

感情スコアP = (幸福・愛情, 嬉しい・楽しい, 安らぎ・信頼, 昂り・興味, 希望)<sup>T</sup>

感情スコアN = (不安・恐れ, 怒り・苛立ち, 哀しみ, 嫌い・不愉快, 退屈・うんざり)<sup>T</sup>

感情スコアPはポジティブな感情、感情スコアNはネガティブな感情をベクトル化したもの

上記予測モデルで過去の視聴率データ実績に当てはめてみたところ、誤差±1ポイント以内に収まるモデルとなり、一般的な重回帰分析に比べ、精度がUPしました。



こちらでいよいよ、今期春ドラマの最終回視聴率予測に挑みたいと思います！が、その前に、最終回視聴率予測の答え合わせまで少し間があるのと、過去の視聴率推移の傾向から、5週目～8週目の視聴率推移が重要ということが分かっていますので、こちらの予測モデルで最終回視聴率とあわせて、前段で直近9週目視聴率の予測も行ってみようと思います。

その結果…

2017年4月期春ドラマ 最終回視聴率予測結果										
放送局	放送曜日	放送時間	ドラマタイトル	ジャンル	主演	実績視聴率			予測視聴率	
						初回	2週目	3週目	9週目	最終回
フジ	月	21:00	貴族探偵	ミステリー	相葉雅紀	11.8%	8.3%	9.1%	8.8%	10.7%
フジ	火	21:00	CRISIS 公安機動捜査隊特捜班	刑事	小栗旬、西島秀俊	13.9%	11.2%	12.0%	11.9%	13.2%
フジ	火	22:00	あなたのことはそれほど	恋愛	波瑠、東出昌大	11.1%	9.0%	9.5%	12.1%	14.3%
テレ朝	水	21:00	警視庁捜査一課9係 season12	刑事	井ノ原快彦	14.5%	9.6%	10.8%	12.6%	13.0%
日テレ	水	22:00	母になる	ヒューマン	沢尻エリカ	10.6%	10.7%	9.3%	10.3%	12.2%
テレ朝	木	21:00	緊急取調室 (第2シリーズ)	刑事	天海祐希	17.9%	14.2%	12.4%	14.5%	17.4%
フジ	木	22:00	人は見た目が100パーセント	ビジネス	桐谷美玲	9.5%	6.4%	6.0%	6.6%	7.6%
TBS	金	22:00	リバース	サスペンス	藤原竜也	10.3%	6.3%	10.5%	9.5%	10.6%
日テレ	土	22:00	ボク、運命の人です。	恋愛	亀梨和也	12.0%	9.6%	9.2%	8.7%	10.5%
TBS	日	21:00	小さな巨人	刑事	長谷川博己	13.7%	13.0%	11.7%	13.3%	14.0%
フジ	日	21:00	櫻子さんの足下には死体が埋まっている	ミステリー	観月ありさ	6.9%	6.5%	5.7%	5.7%	6.6%
日テレ	日	22:00	フランケンシュタインの恋	サスペンス	綾野剛	11.2%	7.3%	8.4%	9.6%	9.6%

※2017/6/2時点予測

さて、はたして結果はどうなるでしょうか…?

最後に

ドラマ視聴率については、ドラマが始まる前、主演の発表辺りから、頻繁にその「予想」がなされ、盛り上がります。前評判や期待度が高いドラマでも蓋を開けてみたらまいまい盛り上がらない…、なぜ視聴率を稼げなかったのか、あらゆる方面から考察され、話題にされ、私たちの目や耳に飛び込んできます。

ソケット視聴率予測シリーズ第1回目の本レポートでは、3週目までの視聴率実績を元に、MSDBから抽出したあらすじデータを元に感情スコア値を出し視聴率予測モデルを構築しましたが、前述で少し触れたとおり、あらすじデータ次第で、算出される感情スコア値が変動し、それが予測精度にも直結しますので、MSDBでのあらすじデータ自体の質の向上、また主演俳優単体の評価だけではなく、キャスト全体を通して人に与えるイメージや相乗効果ももたらす影響、そして、Tweet数のみならず、Tweetコメントからの感情スコア値算出とその相関反映など、人の感情をキーとした本格的な予測モデルへのチューニングに向けて、まだまだ改善の余地があり、そこに可能性が広がっていると考えています。今回は、9週目視聴率予測の答え合わせとプロトタイプ版予測モデルのチューニング、あわせて、今回MSDBから抽出したメタ情報、付帯情報と、視聴率推移とで時期ごとの傾向がみられた相関などをご紹介できればと思います。

ソケットが目指す未来は、もちろん視聴率予測自体にとどまりません。

より細かい各話あらすじデータの感情分析、映像構成、セリフのひとつひとつから、人の感情がどのように突き動かされ、どのように受け止められ、惹きつけられるのか、音楽同様に人の感情を理解するコンテンツとして、ソケットの映像研究も今後加速していきます。

ソケットの感情分析エンジンがそれを理解することで、その時代に求められるドラマの本質、キャスト、構成がどのようなものか、ヒット予測に伴う制作支援でもバリューを提供していけるものと思います。

過去の事実情報からだけでは予期できないヒット要因を、人の感情をキーとして予測する、ソケットの挑戦にご期待ください。

「ソケットレポート」の記事のご利用、また共同研究などに関するご相談はこちらのお問い合わせフォームよりご連絡ください。

<http://www.sockets.co.jp/inquiry/contact/policy.html>

●公開日

平成 29 年 6 月 2 日（金）

●ソケットレポート

2017 年春ドラマ視聴率を予測する シリーズ第 1 回

[http://www.sockets.co.jp/kansei/kansei\\_report03.html](http://www.sockets.co.jp/kansei/kansei_report03.html)

**株式会社ソケット**：<http://www.sockets.co.jp/>

株式会社ソケットは、平成 12 年 6 月に設立された「人の気持ちをつなぐ」という事業目的を持ったデータベースサービス会社です。現在、KDD I 株式会社、株式会社 NTT ドコモ、カルチュア・コンビニエンス・クラブ株式会社などへの音楽や映像、書籍などを中心とした検索・レコメンド・ストリーミング・データ提供・アナリティクスなどのデータ関連サービスを行っております。

**本リリースに関するお問い合わせ先：**

株式会社ソケット

〒151-0051 東京都渋谷区千駄ヶ谷 4-23-5 JPR 千駄ヶ谷ビル 3F

担当：コーポレート本部 小田嶋

Tel：03-5785-5518 Fax：03-5785-5517

E-mail：ir@sockets.co.jp

以上