

平成 28 年 12 月 28 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、紅白歌合戦の曲目の歌詞データから “感情”を分析、さらに今年の勝敗を予測

---

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、熟練スタッフが実際に商品を「見て」「聞いて」「読んで」約 2,000 項目に及ぶ感覚表現メタをチョイスしたオリジナルの人力メタデータである「感性メタ（※）」を教師データとして、2 部制がはじまった第 40 回 1989（平成元年）以降から、本年度第 67 回 2016（平成 28 年）の紅白歌合戦で、紅白それぞれに分かれて披露された楽曲を対象に歌詞データから“感情”を分析、さらにその結果から今年の勝敗予測を公開いたしました。

（※）感性メタとは、特に日本の繊細な感性表現にこだわり、たとえば音楽の歌詞においては全てを読み、内容の真意を理解してから付与するなど機械には出来ない膨大な作業によって生まれた国内唯一のエンターテインメントメタ(ゲノム)データです。

---

### 紅白歌合戦の歌詞データの感情をスコアリング

---

今回は、2 部制がはじまった、第 40 回 1989（平成元年）以降から、先日 2016 年 12 月 19 日に曲目が発表された本年度第 67 回 2016（平成 28 年）の紅白歌合戦で、紅白それぞれに分かれて披露された楽曲を対象とすることにしました。

対象期間：第 40 回 1989（平成元年）～第 67 回 2016（平成 28 年）

※紅組・白組の垣根を越えて披露された楽曲は対象外とする

※『紅白エディション』はオリジナル楽曲の歌詞を解析対象とする

※メドレーは対象外とする

全 28 回、全対象楽曲 1,358 曲　うち紅組 692 曲、白組 666 曲

今回のメドレーなどの対象外含む全披露楽曲 1,464　対象率 93%

そして、今回の分析では、感情に関するキーワードを以下の 10 種類に絞りました。

|       |       |        |        |       |
|-------|-------|--------|--------|-------|
| 幸福・恋愛 | 好き・喜び | 安らぎ    | 哀しみ    | 昂ぶり   |
| 希望    | 不安・恐怖 | 怒り・苛立ち | 嫌い・不愉快 | もどかしさ |

まずは、各楽曲の歌詞データに現れるキーワードの中から、感情に関するフレーズを抽出し、上記 10 種ごとに分類します。この分類の際にソケットの感性メタを教師データとして分類しています。

事例：第 67 回 2016（平成 28 年）白組出場 RADWIMPS／前前前世

| no | sentence                              | emotion_words                          |
|----|---------------------------------------|----------------------------------------|
| 1  | やっと 眼を 覚ましたかい それなのに なぜ 眼も 合わせやしないん だい |                                        |
| 2  | 遅いよと 怒る 君 これでも やれるだけ 飛ばして きたんだよ       | #好き・喜び:やれる #哀しみ:遅い #昂ぶり:飛ばす #怒り・苛立ち:怒る |
| 3  | 心が 身体を 追い越して きたんだよ                    |                                        |
| 4  | 君の 髪や 瞳だけで 胸が 痛いよ                     | #幸福・恋愛:髪 瞳 胸 #嫌い・不愉快:胸 痛い              |
| 5  | 同じ時を 吸いこんで 離したくないよ                    | #哀しみ:離したくない                            |
| 6  | 遥か 昔から 知るその声に                         |                                        |
| 7  | 生まれて はじめて 何を 言えばいい                    | #哀しみ:言え #希望:言う                         |
| 8  | 君の 前 前 前世から 僕は 君を 探し はじめたよ            |                                        |
| 9  | そのぶきっちょな 笑い方を めがけて やってきたんだよ           | #好き・喜び:笑い                              |
| 10 | 君が 全然 全部 なくなって チリチリに なったって            | #哀しみ:なくなっ                              |
| 11 | もう 迷わない また 1 から 探し はじめるさ              | #哀しみ:探す                                |
| 12 | むしろ 0 から また 宇宙を はじめて みようか             | #希望:みる 宇宙                              |
| 13 | どっから 話すかな 君が 眠って いた 間の ストーリー          | #幸福・恋愛:間 #哀しみ:眠っ #昂ぶり:ストーリー            |

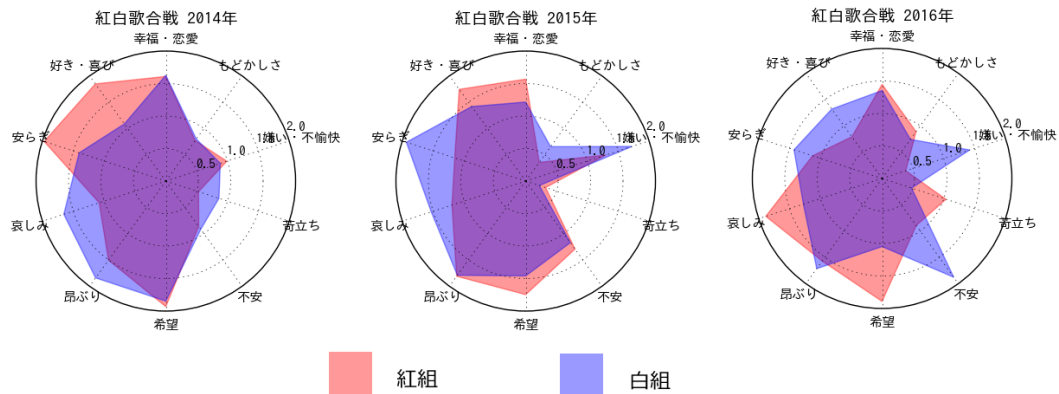
感情ごとに分類したフレーズをスコア化します。

| no | sentence         | 幸福・恋愛     | 好き・喜び     | 安らぎ | 哀しみ       | 昂ぶり       | 希望        | 不安・恐怖 | 怒り・苛立ち    | 嫌い・不愉快    | もどかしさ |
|----|------------------|-----------|-----------|-----|-----------|-----------|-----------|-------|-----------|-----------|-------|
| 1  | やっと 眼を 覚まし...    | 0         | 0         | 0   | 0         | 0         | 0         | 0     | 0         | 0         | 0     |
| 2  | 遅いよと 怒る 君 こ...   | 0         | 0.6       | 0   | 0.7727273 | 0.6055363 | 0         | 0     | 0.9245283 | 0         | 0     |
| 3  | 心が 身体を 追い越し...   | 0         | 0         | 0   | 0         | 0         | 0         | 0     | 0         | 0         | 0     |
| 4  | 君の 髪や 瞳だけで...    | 2.4959381 | 0         | 0   | 0         | 0         | 0         | 0     | 0         | 0.7802469 | 0     |
| 5  | 同じ時を 吸いこんで...    | 0         | 0         | 0   | 0.6315789 | 0         | 0         | 0     | 0         | 0         | 0     |
| 6  | 遥か 昔から 知るその...   | 0         | 0         | 0   | 0         | 0         | 0         | 0     | 0         | 0         | 0     |
| 7  | 生まれて はじめて...     | 0         | 0         | 0   | 0.7837838 | 0         | 0.7692308 | 0     | 0         | 0         | 0     |
| 8  | 君の 前 前 前世から...   | 0         | 0         | 0   | 0         | 0         | 0         | 0     | 0         | 0         | 0     |
| 9  | そのぶきっちょな...      | 0         | 0.9810479 | 0   | 0         | 0         | 0         | 0     | 0         | 0         | 0     |
| 10 | 君が 全然 全部 なく...   | 0         | 0         | 0   | 0.6842105 | 0         | 0         | 0     | 0         | 0         | 0     |
| 11 | もう 迷わない また...    | 0         | 0         | 0   | 0.6422018 | 0         | 0         | 0     | 0         | 0         | 0     |
| 12 | むしろ 0 から また 宇... | 0         | 0         | 0   | 0         | 0         | 2.0030874 | 0     | 0         | 0         | 0     |
| 13 | どっから 話すかな 君...   | 1.0810811 | 0         | 0   | 0.75      | 0.6       | 0         | 0     | 0         | 0         | 0     |

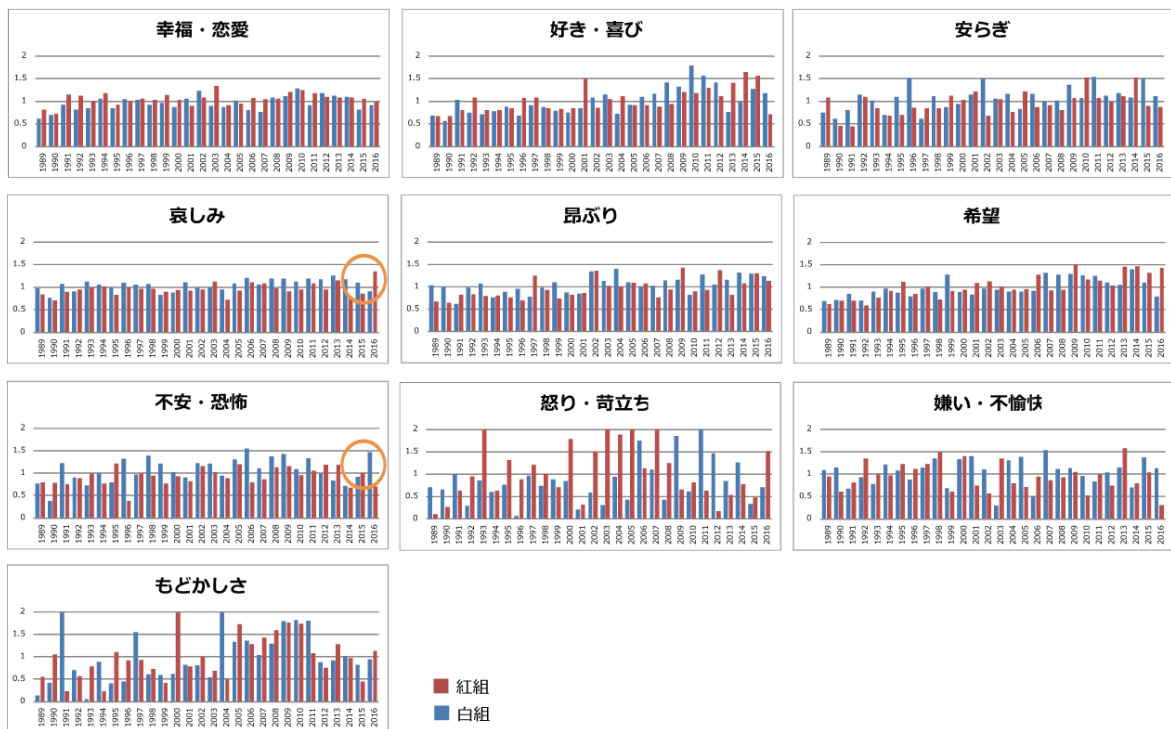
全対象楽曲 1,358 曲のスコア化が完了したら、年代・紅白の組ごとに数値を集計、個々の感情スコア値ごとに正規化します。

ちなみに、その結果を、年ごとに最大値が 2 になるように平準化しグラフに出力したものがこちらです。

※下記は、本年度 2016 年分を含んだ直近 3 年間のグラフ



さらに、感情別に各年代・紅白別の推移をグラフで見てみましょう。



今年 2016 年の楽曲は、紅組は「哀しみ」のスコアが高く、白組は「不安・恐怖」のスコア

が高いようです。

それぞれどの楽曲が影響しているのか見てみると、

| 紅組「哀しみ」の高スコア楽曲    | 白組「不安・恐怖」の高スコア楽曲            |
|-------------------|-----------------------------|
| 1. 夜桜お七／坂元冬実      | 1. 硝子の少年／KinKi Kids         |
| 2. SAKURA／いきものがかり | 2. PERFECT HUMAN／RADIO FISH |
| 3. ごめんね…／高橋真梨子    | 3. ハリケーン・リリ、ポストン・マリ／AAA     |

どうでしょうか。それぞれの歌詞が思い浮かぶ方は、なんとなく高スコアの要因がイメージできるのではないのでしょうか。

ちなみに、感情に関わらず歌詞によく登場するフレーズは特に年ごとに特徴があるわけではなく、一般的に J-POP でよくつかわれるフレーズでした。

| 全28回紅白楽曲合計 |                             | 出現回数 | 2016年紅組楽曲         | 出現回数 | 2016年白組楽曲     | 出現回数 |
|------------|-----------------------------|------|-------------------|------|---------------|------|
| 1.         | 夢、夢見る、夢ある、夢追う、夢描く 夢追いかける    | 943  | 抱きしめ、抱きしめる、抱き合う   | 17   | 吹く            | 5    |
| 2.         | 愛、愛し、愛す、愛する、愛しい             | 686  | 愛、愛し、愛する          | 12   | 流れ、流れる、流れて    | 4    |
| 3.         | 胸、胸抱く、胸に刻む、胸を張る、胸が痛い、胸に残る   | 622  | 生き、生きる            | 6    | 消えて、消えない、消えぬ  | 4    |
| 4.         | 涙、涙流す、涙こぼれる、涙ぐむ、涙落ちる、涙あふれる  | 609  | 忘れて、忘れないよ、忘れはしない  | 5    | 輝く            | 3    |
| 5.         | 夜、今夜、夜明け、夜空、夜更け、一夜、真夜中      | 541  | 歩いて、歩いた           | 4    | 探してる、探しに      | 3    |
| 6.         | 心、心決める、心から、心躍る、心ときめく        | 531  | やってくる             | 3    | 光る            | 3    |
| 7.         | 時、時間、時代、時計、時は流れ、時が過ぎる       | 518  | 逢いたい、めぐり逢い        | 3    | 揺れる           | 3    |
| 8.         | 生き、生きる、生きよ、生きろ              | 508  | 泣いて               | 3    | 超えてゆけ         | 3    |
| 9.         | 恋、恋し、恋しい、恋人、恋しく、恋する、恋愛、恋に落ち | 500  | 消えそう、消えていく、消えない   | 3    | 思う、思うほどに、思うんだ | 3    |
| 10.        | 信じ、信じる、信じあえる、信じよ            | 349  | 置いていけ、置いてけ、置いていこう | 3    | 照らす           | 2    |

---

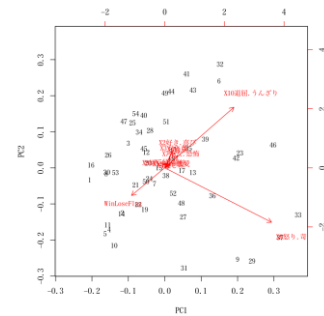
感情スコアを分析し、勝敗の予測モデルを作ってみる

---

さて、紅白の勝敗といえば、楽曲そのものでの勝負というわけではなく、対抗形式で歌や演奏を披露し、全体のショーを通して最後に勝敗を決める、というものであり、歌詞の感情分析と勝敗の因果関係と結び付けるのは難しいものがあると思いますが、そこをあえて、今回は、この感情分析と過去の勝敗のみで他の相関・要素を入れずにシンプルに分析していくことにします。

上述の感情スコア値のグラフを見ても分かる通り、感情スコアと勝敗には、パッとみる限り、相関は特に見当たりませんが、まずは主成分分析などをはじめ「データをよく見る」ことから始めていきます。

そして、今回は勝敗を予測するために、ロジスティック回帰分析で予測モデルを作ってみたいと思います。



ロジスティック回帰分析とは、ある現象の発生する確率を、その現象の起因を説明するために観測された変数群によって説明をするための分析手法です。今回の目的では、勝ち（1.0）か負け（0.0）かを、その現象の規制を説明するために観測された変数群として感情スコアを利用して予測するものです。

このとき 勝ち（1）／負け（0） という値で得られている変数は 1 もしくは 0 という 2 値のいずれかで示され、この 1 もしくは 0 という値を、他の説明変数が決定付けている、もしくは影響を与えているとみなして分析を行ない、数式を構築します。

年別、紅白別の 10 種の感情スコアから分析した回帰係数は下記のようにになりました。

| 感情項目        | 回帰係数         |
|-------------|--------------|
| (Intercept) | -0.165634023 |
| X1幸福,恋愛     | 0.21316932   |
| X2好き,喜び     | 0.19112587   |
| X3安らぎ       | -0.362814407 |
| X4哀しみ       | 1.025946027  |
| X5昂ぶり       | 0.275976328  |
| X6希望        | -0.336567294 |
| X7不安,恐怖     | 0.410326195  |
| X8怒り,苛立ち    | -0.172823138 |
| X9嫌い,不愉快    | -0.302512827 |
| X10退屈,うんざり  | -0.284878219 |

この係数が大きいほど、勝敗への影響度が高いとみなされます。この推定されたモデル式を用いて、過去計 27 年間×紅白別の全 54 の予測値を算出し、年ごとに紅白で比較をして高い予測確率である方が「勝利する」という予測をしました。

---

## 勝敗の予測モデルの正解率と今年の勝敗予想

---

さて、できあがった予測モデルに当てはめた時の、勝敗実績の正解率はどうのような感じでしょうか。

| 予測モデルによる予想 | 予想件数 | ○正解 | ×はずれ | 正解率  |
|------------|------|-----|------|------|
| 紅が勝つと予想    | 9    | 8   | 1    | 89%  |
| 白が勝つと予想    | 17   | 17  | 0    | 100% |
| 対象期間合計     | 26   | 25  | 1    | 96%  |

今回対象とした過去 27 回のうち、予測モデルに当てはめたときに、

・ 紅が勝つ予測：9 件

○紅が実際に優勝：8 件 ×白が実際に優勝：1 件 ⇒正解率 89%

・ 白が勝つ予測：17 件

○白が実際に優勝：17 件 ×紅が実際に優勝：0 件 ⇒正解率 100%

全 27 回分トータルで、正解率 96%という結果になりました。

実際のビッグデータ分析・予測においては、今回のロジスティック回帰分析にとどまらず、ニューラルネットワークでさらに精度を追求、また、主成分分析などを行ったりしますが、分かりづらくなるので、今回はここでとどめて、この予測モデルに当てはめて、今年 2016 年の紅白の勝敗を予測すると…

| 2016年 | 理論値         |
|-------|-------------|
| 紅組    | 0.596766667 |
| 白組    | 0.418666667 |

予測モデルの理論値において、紅白で比較して高い値の方が「勝利する」ということは…ソкетов感情分析からの今年 2016 年の紅白歌合戦は、ずばり紅組の優勝が予想されました。さて、結果はいかに…？！

最後に

今回は紅白歌合戦の歌詞を感情分析して、さらに勝敗予想してみるというレポートでしたが、「感情分析エンジン」の技術開発では、様々な角度から研究を行っています。

様々なデバイスを利用しているユーザーの感情を読み取る場合、単純にユーザーアクションやコメント内容からのみの感情抽出だけでは人の複雑な感情を推測はできません。

感情状態の時系列分析をおこない、状態変化を推定するような処理が必要となります。

人間と同等の知能をもつ人工知能に、人の感情を理解できる『感性』をもたせられたら…

人の『感性』や『感情』を扱ってきたソкетовだから目指すべきこと、そして実現できること。

今後もソкетовの技術研究開発にご期待ください。

「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 28 年 12 月 28 日（水）

●ソケットレポート

[http://www.sockets.co.jp/kansei/kansei\\_report01.html](http://www.sockets.co.jp/kansei/kansei_report01.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モバイルインターネット分野を中心としたサービス、アプリケーション、データベースの開発会社です。現在、KDD I 株式会社や株式会社NTTドコモなどへの音楽や映像、電子書籍などの検索サービスやストリーミングサービスに加え、カルチュア・コンビニエンス・クラブ株式会社や楽天株式会社への作品おすすめ（レコメンド）機能の提供、国内最大級の楽曲データベースを活かした自社開発パーソナライズドサービス「LIFE's radio」の提供などを行っております。

**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](mailto:ir@sockets.co.jp)

以上