

動画コンテンツにおける 発話単語と高評価数の関連分析

1 立命館大学情報理工学部

2 立命館グローバル・イノベーション研究機構

瀧岡 要¹、安尾 萌²、Shan Junjie¹、○西原 陽子¹

研究の動機

- 動画共有サイトに投稿される多くの動画、
内容を見る前に面白いものかどうかを知りたい
- アップロードされた瞬間に、それが面白いものかどうかを知りたい
- 動画の絞り込み手段としては、推薦アルゴリズムやメタデータがある
 - 関連する動画を推薦するアルゴリズム[Solaiman2023, Xu2018]
 - 動画作成者や高評価数、低評価数、視聴回数などのメタデータ
- アップロードされた直後の動画に対して、
それが面白いかどうかを評価する手法が少ない

見るまで中身が面白いかがわからない

【最新の動画はこちら】 ▶ すべて再生

37万 回視聴・16 時間前

35万 回視聴・2 日前

18万 回視聴・3 日前

33万 回視聴・5 日前

48万 回視聴・6 日前

47万 回視聴・7 日前

- ・ タイトル, 動画作成者, 視聴回数, アップロード日から判断することが多い
- ・ 1つ見て, 面白くなかったら, 次のを探す, ザッピングする.
- ・ 時間が削られて, もったいない

研究の目的

- 動画が高評価となるかどうかを推定するための**基礎調査**として**動画の高評価数と動画内の発話単語の関連**进行分析
 - 本論文では、面白さの指標として高評価数を用いる
- 分析結果を用いることで、アップロードされた直後の動画であってもその内容から獲得される評価数を推定する手法を提案する
- 新たな動画推薦や動画作成者へのサポート [Samonte2022] に応用する

本研究の仮説

- 動画は複数のメディア（発話情報，視覚情報，音響情報）で構成され，その利用割合は等しくない
 - 発話情報を主とする：解説や実況
 - 視覚，音響情報を主とする：リアクション動画やクリップ集
- 仮説1：動画が主として発話を情報伝達に用いるならば，発話に含まれる語彙数と動画の高評価数は関連する
- 仮説2：動画が主として視覚と音響を情報伝達に用いるならば，発話に含まれる語彙数と動画の高評価数は関連しない

分析手順の概要

- 1. 動画データ, メタデータの収集
- 2. 発話を書き起こしたテキストの作成
- 3. テキスト内の単語の統計情報の獲得
- 4. 高評価数と単語の統計情報の相関係数算出

1. 動画データ，メタデータの収集 (1)

- 主として発話を情報伝達に用いる動画と，主として視覚，音響情報を情報伝達に用いる動画を収集した
- 各テーマで10件ずつ，動画を収集した

発話を主とする動画のテーマ	視覚，音響を主とする動画のテーマ
1. きさらぎ駅の解説	6. 事故の衝撃映像
2. 月見バーガーの商品レビュー	7. 猫の可愛い動画
3. Getting Over It with Bennett Foddyのゲーム実況	8. Apex Legendsのクリップ集
4. 無印良品の商品のみを使ったメイク動画	9. メントスコーラのリアクション動画
5. たぬきの生態解説	10. カルボナーラのレシピ動画

1. 動画データ, メタデータの収集 (2)

- 収集時の条件：
 - 公開されてから1年以上が経過しているもの
 - 長さは10分程度
 - 評価数が少ないものから多いものまでを満遍なく
 - あとで相関係数を評価したいため

高評価数の統計データ

動画テーマの ID	最小値	最大値	平均値
1	2	11,736	1,680
2	0	2,857	727
3	0	18,564	5,559
4	153	31,045	4,343
5	8	11,806	1,952
6	8	4,801	1,210
7	5	86,600	11,880
8	12	4,169	1,089
9	0	27,704	8,350
10	13	19,292	5,962

2. 発話を書き起こしたテキストの作成

- 収集した動画の発話を書き起こしたテキストを作成する
- OpenAIのWhisperのlargeモデルを用いて書き起こし、その後人手でチェックを行なって完成させた
 - 動画内で使用されているBGMの歌詞やゲームキャラクターの発話を書き起こされる場合があるため

書き起こされたテキストの例

そこに着いたら二度と帰れない。
電車が止まったのは存在しない駅。
帰ろうとするとまるで追いかけるようになり響く太鼓の音。
その駅の名はきさらぎ駅。
そんな伝説の怪談きさらぎ駅を民族学的に解剖しよう。
こんみこ、ようこそ民族学研究室へ。
主任研究員の天道みこです。
今回は誰もが聞いたことがあるであろう伝説の怪談実況、きさらぎ駅。
こちらを民族学の視点から考察した記事を見つけたので紹介しようと思う。
結論から言ってしまうと主人公は儀式に招かれたのでは？

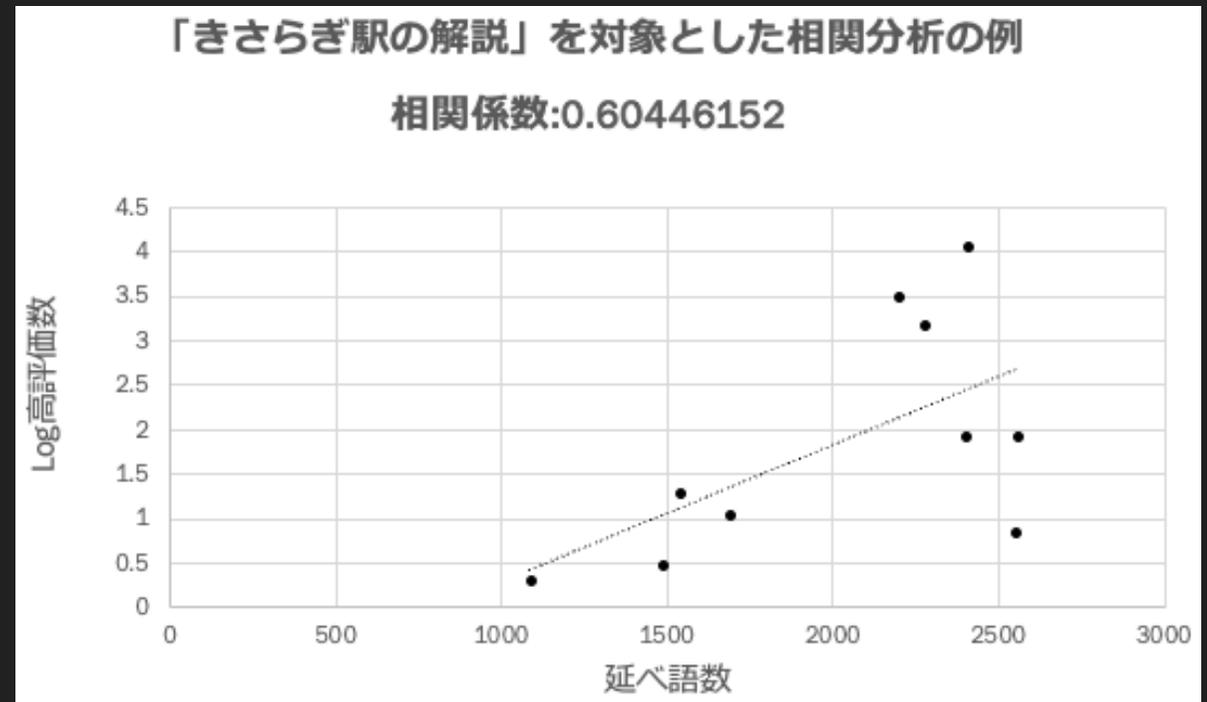
【都市伝説民俗学】人気第1位“きさらぎ駅”を民俗学的に徹底考察！【説天道巴弧Vtuber】より

3. テキスト内の単語の統計情報の獲得

- 書き起こされたテキストに対し、MeCabとNEologdで形態素解析を行い、単語の統計情報を得る
- 使用する品詞は12種類：
 - 記号, 助詞, 助動詞, 接続詞, 接頭詞, 名詞, 動詞, 連体詞, 感動詞, 形容詞, フィラー, 副詞
- 品詞ごとの(1) 延べ語数, (2) 異なり語数, (3) 延べ語数/秒, (4) 異なり語数/秒を各動画に対して出す
 - 1秒あたりの数値を出すのは, 動画の長さがそれぞれ異なっているから

4. 高評価数と単語の統計情報の相関係数算出

- テーマごとに10本の動画について、高評価数と単語の統計情報（ $12 \times 4 = 48$ 種類）の相関係数を算出する
- 相関係数が0.6以上で相関があり、0.3以上0.6未満で中程度の相関があるとする



分析結果1：発話情報を主とする動画について(1)

- 発話情報を主とする動画で高評価数と相関が見られたものは、**形容詞**に関する指標であった
 - 中程度の相関が見られたものは、自立語（名詞，動詞，副詞，連体詞），付属語（助詞，助動詞）それぞれあった。
- 異なり語数よりも**延べ語数**と相関が見られることが多かった
 - 1秒あたりの数値と相関が見られることが多かった

相関係数が ^s 0.6 以上	
指標	テーマ数
形容詞 (延べ語数)	2
形容詞 (延べ語数)/s	2
相関係数が ^s 0.3 以上 0.6 未満	
指標	テーマ数
名詞 (異なり語数)/s	3
名詞 (延べ語数)/s	3
助詞 (延べ語数)/s	3
助動詞 (延べ語数)/s	3
動詞 (延べ語数)/s	3
副詞 (異なり語数)/s	3
連体詞 (延べ語数)/s	3

分析結果1：発話情報を主とする動画について(2)

- 形容詞に関する指標で相関が見られた理由は、
動画作成者の主観的な評価が形容詞に現れることが多く、それが高評価数と関連した
- 仮説1は支持された

発言例

こうやってたまに食べると**美味しい**っすよねー
ソースが**甘い**

また、マレビット信仰には**新しい**DNA
を村にもたらすという側面もある

怪しい状況になり投稿は途絶えてしまう
意外と操作**難しい**んだよ

危ないまって

ねえ、見てめっちゃチーク**可愛い**の
薄いけどちゃんと目が大きく見える

日本人からすれば**珍しい**とは言えない動物だと思うよ
いや、たぬきは**幅広い**体のせいか

分析結果2：視覚，音響情報を主とする動画について(1)

- 視覚，音響情報を主とする動画で，高評価数と相関が見られたものはなかった
- 中程度の相関が見られたものは，フィラーや感動詞，付属語（助詞，助動詞）であった
- 異なり語数と相関が見られることが多かった
 - 1秒あたりの数値と相関が見られることが多かった

相関係数が 0.3 以上 0.6 未満	
指標	テーマ数
フィラー (延べ語数)/s	3
助詞 (異なり語数)	2
助詞 (異なり語数)/s	2
連体詞 (異なり語数)/s	2
感動詞 (異なり語数)	2
感動詞 (異なり語数)/s	2

分析結果2：視覚，音響情報を主とする動画について(2)

- フィラーや感動詞に関する指標で中程度の相関が見られた理由は，動画作成者の視覚情報に対する感想がフィラーや感動詞に現れることが多かったからと考えられる
- フィラーや感動詞は語彙としないとするならば，仮説2は支持された

発言例

うわ、結構多いんだな

え、今ぶつかったよね

あー、こらこらこらこら

なんかしっぽは

じゃーあの一

え、つっよ

おい、メントスコーラに勝っちゃったぞ、こいつ

なんか嫌やね

いや、ほんと

えーと、ローマとか

本発表のまとめ

- 動画が高評価となるかどうかを推定するための基礎調査として動画の高評価数と動画内の発話単語の関連を分析した
- 分析の結果、以下の2つが分かった
- (1) 動画が主として発話を情報伝達に用いるならば、発話に含まれる語彙数（特に形容詞）と動画の高評価数は関連する
- (2) 動画が主として視覚と音響を情報伝達に用いるならば、発話に含まれる語彙数と動画の高評価数は関連しない
 - フィラーや感動詞と中程度の相関は見られた