

テレビドラマに関する関心动向・感想のウェブマイニングにおけるインパクトの分析

Analyzing Impressiveness in Web Mining of Concerns and Reviews on TV Drama

藤田 拓也[†] 久保 遼馬[†] 川口 輝太[†] 前田 竜治[†] 宇津呂武仁[†]
小林 彰夫^{††} 西崎 博光^{†††} 河田 容英^{††††}

[†] 筑波大学 大学院システム情報工学研究科 〒 305-8573 茨城県つくば市天王台 1-1-1

^{††} 筑波技術大学産業技術学部 〒 305-8520 茨城県つくば市天久保 4-3-15

^{†††} 山梨大学大学院総合研究部 〒 400-8511 山梨県甲府市武田 4-3-11

^{††††} (株) ログワークス 〒 151-0053 東京都渋谷区代々木 1-3-15 天翔代々木ビル 6F

あらまし 本論文では、テレビドラマ視聴者がドラマ視聴後にウェブ上で行うドラマ関連関心动向・感想・レビュー類の情報探索過程を支援することを目的として、ブログ・ドラマ関連サイト・ツイッター等のウェブページからの情報収集・集約を行うウェブマイニング技術を提案する。本論文では、特に、ブログ・ドラマ関連サイトにおける記事タイトルにおけるインパクトの判定手法について述べる。

キーワード テレビドラマ, 関心动向, 感想, ブログ, インパクト

1 はじめに

本論文では、テレビドラマ視聴者がドラマ視聴後にウェブ上で行うドラマに関する関心・感想・レビュー類の情報探索過程を支援することを目的として、ドラマに関するブログや情報サイト等のウェブページからの情報収集・集約を行うウェブマイニング技術を提案する。本論文でマイニング対象とするウェブページは、ドラマ視聴者によるドラマに関する関心や感想を含むブログ(図1)、及び、ドラマ関連グッズを販売する目的のECサイト、エンターテインメント系のネットニュースサイト、見逃し配信サービスを紹介するサイト等、様々なウェブページ(図2)が含まれる。本論文においてそれらのウェブページ群をマイニングして、情報収集・集約を行った結果を提供する利用者像としては、各ドラマの熱狂的な視聴者が挙げられるが、そのような熱狂的な視聴者に限らず、偶然当該ドラマに関心を持ち、情報収集をしようと試みる一時的な視聴者も重要な利用者として位置付けられる。

本論文では、2018年10~12月のクールを放送時期とする4つのドラマ「リーガルV」、「下町ロケット」、「大恋愛」、「僕らは奇跡でできている」を対象とした。上述の4つのドラマに関連する話題が掲載されている可能性の高いウェブページを網羅的に収集し、各ドラマに関連するか否かの区別、ブログか否かの区別、ブログの場合の感想の有無、等について人手分析を行った¹。その結果、表1に示すように、各ドラマに関して収集さ

れたウェブページ数が10以上のウェブサイトが、各ドラマごとに約300~400程度存在し、そのうち、約40~60サイトは、当該ドラマに関する感想を掲載したブログであった。さらに、それらの当該ドラマに関する感想を掲載するブログとは別に、当該ドラマについて、感想以外の関心事項を掲載するブログが約30~40サイト存在することが確認された。

この分析結果をふまえて、文献[6]では、テレビドラマ視聴者がドラマ視聴後にウェブ上で行うドラマに関する関心・感想・レビュー類の情報探索過程を支援することを目的として、ドラマに関するブログや情報サイト等のウェブページからの情報収集・集約を行うウェブマイニング技術の枠組み(図3)を提案した。この枠組みにおいては、まず、各ドラマに関するブログ、ECサイト、ネットニュース、見逃し配信サービスを紹介するサイト等、様々なウェブページが収集される。そして、「感想・登場人物・キャスト・ドラマ特有のキーワード²・ロケ地や視聴率などのドラマに関する関心事項」の5種類のコンテンツに分類され、冗長な情報の集約が行われる。そして、集約された5種類のコンテンツをドラマ視聴者に提供する。この枠組みの中の要素技術として、文献[6]では、文単位での当該ドラマ関連判定および感想の有無判定を行い、その文単位の結果に基づき、ウェブページ単位での当該ドラマ関連判定および感想の有

1: ブログ以外のウェブページにおいても何らかの感想を含む文が掲載されているが、それらの文において感想が述べられている対象は、当該ドラマの内容その

ものではなく、大半がDVDや関連グッズの販売に関してである。その他、ネットニュースにおいて、キャストである俳優へのインタビュー記事が掲載されており、その俳優の発言中に感想表現が存在する場合なども含まれる。

2: 例えば、ロケットを題材とするドラマである「下町ロケット」の場合には、「ロケットエンジン用バルブシステム」、「小型エンジン」、「佃製作所」等が当該ドラマ特有のキーワードの一例となる。

リーガルVに関するブログ

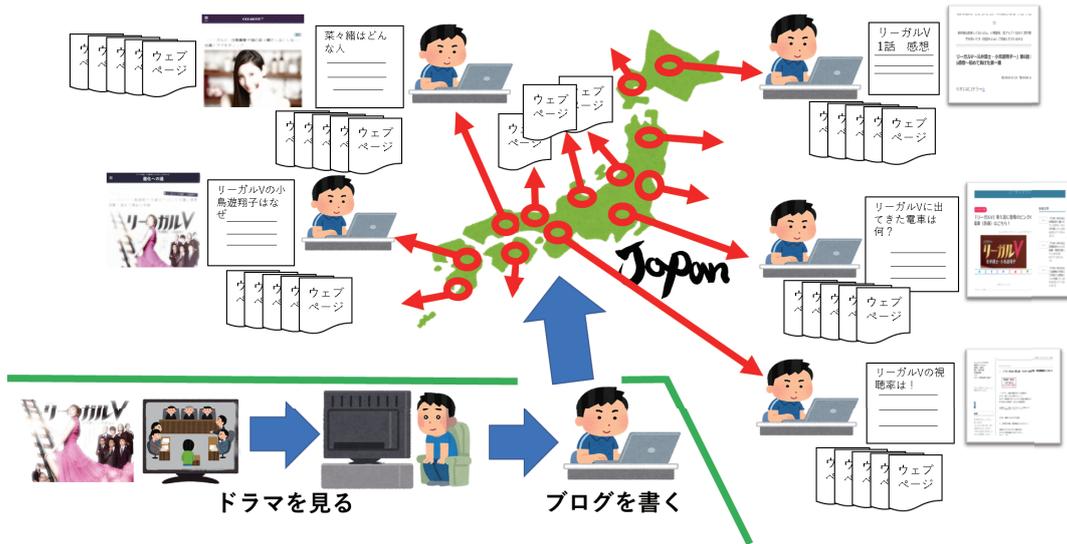


図1 テレビドラマに関する関心・感想についてのブログ記事集合

リーガルVに関するウェブページ

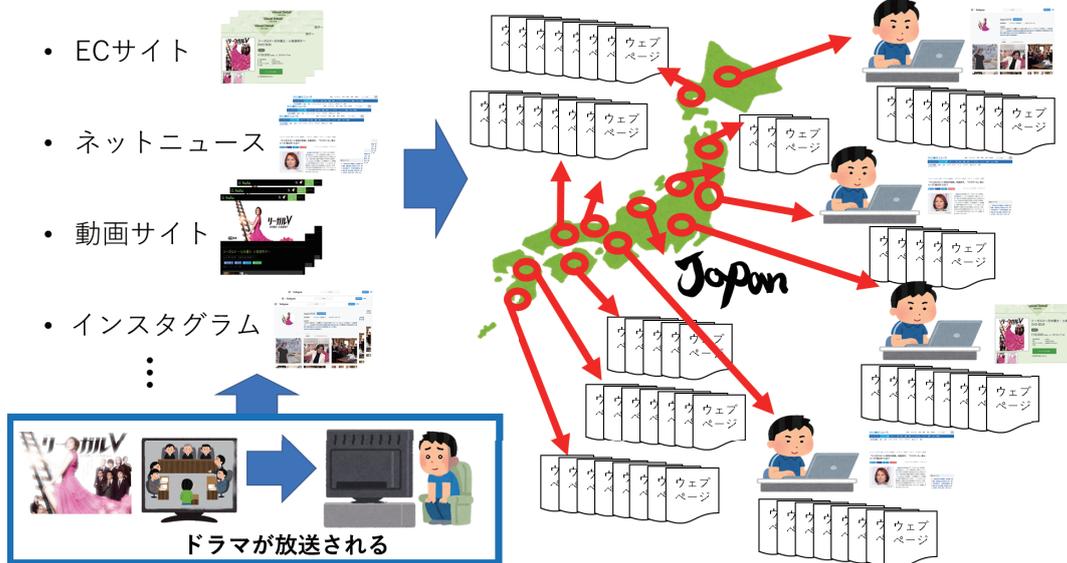


図2 テレビドラマに関連する各種ウェブページ集合 (ブログ以外)

無判定を行う手法、および、その評価結果について述べた。

文献[6]の結果をふまえて、本論文では、文献[6]で訓練・評価対象としたウェブページ群の内容のインパクトの有無について分析を行った結果を述べるとともに、インパクト有無の自動判定手法を提案し評価を行った結果について述べる。具体的には、ブログ・非ブログのウェブページを対象として、まず、ウェブページのタイトルにおけるインパクトの有無と、ウェブページの内容そのものにおけるインパクトの有無の間の相関について調査した。ウェブページタイトルにおけるインパクトの有無を判定する際には、表2に示す判定基準を用いた³。その結果、

調査対象のサンプルの範囲においては、8割以上の事例において、タイトルとウェブページの内容の間にインパクト有無の相関があることが分かった(表3)。このことをふまえて、本論文では、ドラマ名をクエリとしたサジェストによって収集したブログ・非ブログのウェブページを対象として、ウェブページタイトルのインパクトの有無を判定するタスクを設定する。

次に、本論文では、ウェブページタイトルのインパクトの有無判定モデルを訓練するのにあたり、各個人別あるいは嗜好が似通った利用者群ごとに個別の訓練事例を作成し、訓練事例集合別にインパクト有無判定モデルを訓練するのが適切であるの

3: 表2および表3においては、本論文においてインパクト判定作業を行った合計6名の作業による判定結果のうち、後述する kappa 値[2]による一致度が十分に高い作業 A および作業 B を対象として、判定結果を示す。特に、表2

においては、判定基準のうちの大部分が作業 A・作業 B 間で共通となったため、「作業 A・作業 B で共通の判定基準」および「作業 A、または、作業 B のみの判定基準」とに分けて判定基準を示す。

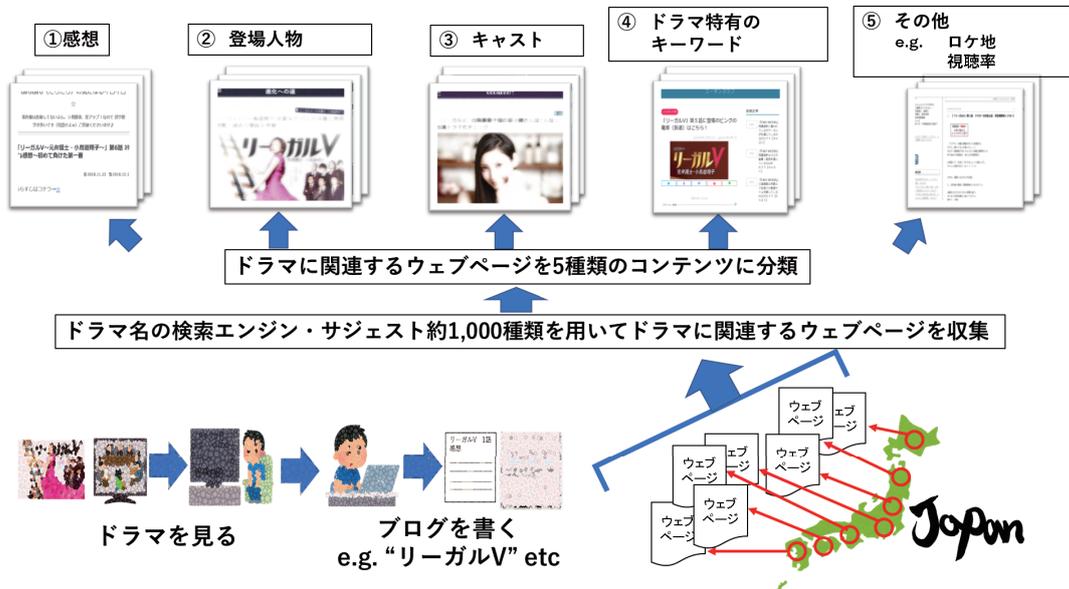


図3 本論文の枠組み: ドラマ関連ウェブページの収集およびドラマに関する関心・感想の分類

表1 各ドラマに対して収集されたウェブサイト・ウェブページの数

ドラマタイトル	サジェスト数	収集された全サイト対象		一ドラマあたりのウェブページ数が10以上のサイト対象						
		ウェブページ数	ウェブサイト数	ウェブページ数	ウェブサイト数				クエリ・フォーカスのドラマに関連しない	
					合計	感想を含むブログ	感想を含まないブログ	ブログ以外		合計
「リーガルV」	672	2,621	791	2,065	338	47	41	218	306	32
「下町ロケット」	782	4,688	1,393	3,287	341	52	35	229	316	25
「大恋愛」	723	3,778	1,244	2,726	401	59	44	251	354	47
「僕らは奇跡でできている」	630	1,912	639	1,457	279	46	34	175	255	24

か、それとも、全利用者に共通の単一インパクト有無判定モデルを訓練するのが適切であるのかを判断するための材料を収集するため、本論文著者を含めた6名を対象として、ウェブページタイトルのインパクトの有無判定結果における kappa 値 [2] の測定を行った⁴。その結果、調査対象の6名のうちの各2名組においては、一致度が低い組が多かったため、インパクトの判定作業の結果においては個人差が大きいことが分かった。この結果をふまえて、本論文では、インパクト判定モデル訓練事例の作成を担当する主たる作業員を選定し、これらの作業員別に訓練・評価事例集合を作成し、各作業員別に異なるインパクト判定モデルを訓練・評価するアプローチを採用する。特に、

4: 調査対象の「下町ロケット」に関する444ページのウェブページタイトルに対して、作業員Aおよび作業員Bを含む6名が、各個人の判定基準に基づきインパクト判定作業を行い kappa 値 [2] の測定を行った。6名のうちの任意の2名の間の kappa 値を測定したところ、作業員A・作業員B間では0.71であったが、他の14組の内訳としては、0.40が1組、残りの13組は0.40より低く、一致度が低い組が多かった。以上の結果より、インパクトの判定作業結果においては個人差が大きいことが判明した。また、表4に示す4つのドラマ全体の4,023ページを対象とするインパクト判定作業結果においても、作業員Aと作業員Bの間の kappa 値は0.65と十分に高い一致度となった。

本論文では、訓練されたインパクト判定モデルを比較することが容易であることを重視して、kappa 値による一致度が十分高い判定作業員Aおよび作業員Bの2人を対象として、各作業員別のインパクト判定モデルの訓練・評価を行う。そして、このタスクに対して、事前訓練済みBERT [3]の分散表現を用いた双方向LSTMを適用し、ウェブページタイトルのインパクト判定が一定の性能で実現できることを示す。

2 ドラマに関連するウェブページのデータセット

各ドラマに対して、表1中の「一ドラマあたりウェブページ数が10以上のサイト対象」欄に該当するウェブページ集合から無作為に約1,000ページ(表4)選定し、各ページのタイトルに対して人手で「インパクトあり」・「インパクトなし」の二値で判定作業を行った。本論文では、kappa 値による一致度が十分高い判定作業員Aおよび作業員Bの2人が判定作業を行うにあたり、各作業員は、それぞれ、「4つのドラマを視聴していない」という立場(作業員A)、および、「4つのドラマを視聴した」という立場(作業員B)で、表2に示す判定基準を用いて判

表2 テレビドラマに関連するウェブページタイトルのインパクト判定基準(主要なもののみ抜粋)

(a) 作業員 A・作業員 B で共通の判定基準

キーワード「視聴率」に言及している
キーワード「ロケ地」に言及している
「極端な評価表現(クズ, 大絶賛など)」が含まれている
「興味を惹く文言(裏事情, 秘密, 農耕アートなど)」が含まれている
「関心・興味がある俳優名」が含まれている
「ドラマ内容に関する専門家の解説」が含まれている

(b) 作業員 A, または, 作業員 B のみの判定基準

「主要な俳優」に言及している
「ストーリーの今後の展開を推測する内容」が含まれている

表3 ウェブページのタイトルと内容の間のインパクトの相関(対象ドラマ:「下町ロケット」)

(a) 作業員 A

		内容のインパクト	
		あり	なし
タイトルのインパクト	あり	88	19
	なし	15	311

(b) 作業員 B

		内容のインパクト	
		あり	なし
タイトルのインパクト	あり	75	17
	なし	7	345

定作業を行った。上記の方法でインパクトの判定を行った結果のデータセットの内訳を表4に示す。この結果において、作業員 A・作業員 B による判定作業結果において、作業員 AB 間で判定結果が一致したページ数は、「リーガルV」において1,003ページ中842ページ、「下町ロケット」において1,000ページ中890ページ、「大恋愛」において1,004ページ中881ページ、「僕らは奇跡でできている」において1,016ページ中920ページであり、いずれのドラマにおいても90%前後のウェブページタイトルにおいてインパクト判定結果が一致していた。

3 ウェブページタイトルのインパクトの分析

本節では、「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の5つの特徴の有無全32通りの組合せにより、ウェブページタイトルのインパクトの分析を行った結果について述べる。本論文では、「下町ロケット」に関するウェブページ1,000ページ(表4)を分析対象とする。具体的には、32通りの各組合せにおいて、各特徴ごとの「インパクトあり」の割合を表5に示す。また、「下町ロケット」における「インパクトあり」クラスのウェブページタイトル集合において、各特徴に該当するタイトルの割合を表6に示す。

表5の分析結果より、ロケ地・視聴率等の情報を含むタイトルは、インパクトありと判定される傾向にあることがわかった。「ドラマ名・評価表現」に関しては、「クズ, 飽きた」等の否定的な表現、「大絶賛, ギャップがすごい」等の肯定的な表現のように極端な評価表現が含まれている場合、インパクトありと判定された。「ドラマ名・主要な俳優」に関しては「インパクトあ

り」の割合が作業員 A・作業員 B ともに低いが、俳優の演技力の分析や撮影裏話の情報等の興味深い情報が含まれる場合にインパクトありとなった。ウェブページのタイトル・内容におけるインパクトの有無の組合わせおよび例を表7に示す。

4 ウェブページタイトルのインパクト判定

本論文の枠組みにおいて、クエリ・フォーカスとして指定されたドラマ名に対して、検索エンジン・サジェストを用いて収集されたウェブページのタイトルにインパクトがあるかを判定するインパクト判定モデルを図4に示す。ウェブページのタイトルにインパクトがあるかを判定するタスクにおいては、事前訓練済みBERT[3]の分散表現を用いた双方向LSTMを適用した。本論文では、PyTorch版⁵BERTを利用し、日本語版BERTの事前訓練済みモデル⁶を用いた。日本語版BERTへの入力文は形態素単位で分割されている必要があるため、日本語形態素解析ツールMeCabのIPA辞書を用いて、形態素単位への分割を行った。双方向LSTMの実装としては、Keras版⁷を利用し、分類のための双方向LSTM層の次元数を128次元とした。

1節で述べたように、本論文では、各作業員別に訓練・評価事例集合を作成し、各作業員別に異なるインパクト判定モデルを訓練・評価するアプローチを採用する。具体的には、表4のデータセットのうち、作業員 A・作業員 B の各々に対して、3つのドラマの事例を訓練事例とし、残りの1つのドラマの事例を評価事例とする4ドラマ交差検定を行い、各作業員ごとのインパクト判定モデルの評価を行う。双方向LSTMモデルにおいて全結合層のsigmoid関数の出力確率の下限値を変化させて、各作業員ごとに再現率・適合率曲線を描いた結果を図5に示す。この結果から分かるように、sigmoid関数の出力確率の下限値が高い範囲では、評価事例割合の小さい「インパクトあり」クラスの方も十分に高い適合率となっている。

本論文では、さらに、4ドラマ交差検定において、片方の作業員の訓練事例で訓練し、もう片方の作業員の評価事例で評価する形で、作業員間を交差する評価も行った。両作業員間では、90%前後のウェブページタイトルにおいてインパクト判定結果が一致し、0.6以上の高いkappa値となっていることから、

5: <https://github.com/huggingface/pytorch-pretrained-BERT>

6: 日本語 Wikipedia の 1,800 万文を用いて事前訓練された Japanese_L-12_H-768_A-12_E-30_BPE.zip モデル (<http://nlp.ist.i.kyoto-u.ac.jp/index.php?BERT> 日本語 Pretrained モデル)

7: Keras(<https://keras.io/>) 上の実装を用いた。

表4 ウェブページタイトルのインパクト判定用訓練・評価データセット

(a) 作業員 A				(b) 作業員 B			
ドラマタイトル	インパクトあり	インパクトなし	合計	ドラマタイトル	インパクトあり	インパクトなし	合計
「リーガル V」	325	678	1,003	「リーガル V」	260	743	1,003
「下町ロケット」	229	771	1,000	「下町ロケット」	183	817	1,000
「大恋愛」	275	729	1,004	「大恋愛」	190	814	1,004
「僕らは奇跡でできている」	204	812	1,016	「僕らは奇跡でできている」	128	888	1,016

表5 「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の組合せにおける「インパクトあり」の割合 (対象ドラマ: 「下町ロケット」)

特徴	作業員 A		作業員 B	
	1,000 ページに対する 出現頻度: インパクト (あり/なし)	「インパクトあり」 の割合	1,000 ページに対する 出現頻度: インパクト (あり/なし)	「インパクトあり」 の割合
ドラマ名・視聴率・ 評価表現	24 / 0	100%	23 / 1	96%
ドラマ名・視聴率	21 / 0	100%	20 / 1	95%
ドラマ名・ロケ地	17 / 1	94%	17 / 1	94%
「ドラマ名・主要な俳優・ ロケ地・視聴率・ 評価表現」の その他の組合せ	21 / 8	72%	13 / 16	45%
ドラマ名・評価表現	47 / 74	39%	44 / 77	36%
ドラマ名・主要な俳優・ 評価表現	23 / 45	34%	17 / 51	25%
ドラマ名・主要な俳優	30 / 116	21%	16 / 130	11%
その他	4 / 22	15%	3 / 23	12%
ドラマ名	42 / 505	8%	30 / 517	5%
合計	229 / 771	23%	183 / 817	18%

表6 インパクトがあるタイトル中に含まれる「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の組合せの割合 (対象ドラマ: 「下町ロケット」, 作業員 A・作業員 B の内訳の平均値の降順)

特徴	作業員	
	A	B
ドラマ名・評価表現	21%	24%
ドラマ名	18%	16%
ドラマ名・視聴率・評価表現	11%	13%
ドラマ名・主要な俳優	13%	9%
ドラマ名・視聴率	9%	11%
ドラマ名・主要な俳優・評価表現	10%	9%
ドラマ名・ロケ地	7%	9%
「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」のその他の組み合わせ	9%	7%
その他	2%	
合計	100%	

作業員間を交差する本評価においても、図5 とほぼ同等の結果となった。

5 関連研究

BERT [3] を用いた関連研究の一例として、レビューの理解や感情の分析において利用するもの [1], [4], 議論の予測において

利用するもの [5], 意見や主張の分析において利用するもの [7], イベントの検出において利用するもの [8] 等が挙げられる。一方、本論文では、テレビドラマ視聴者による情報探索過程支援を目的として、ドラマに関連するウェブページのタイトルのインパクト判定タスクを設定し、このタスクに BERT を適用しており、先行研究とは問題設定が異なる。また、文献 [6] では、当該ドラマ関連判定および感想判定の2つのタスクを設定して

表7 テレビドラマに関連するウェブページのタイトル・内容におけるインパクトの有無の組合せ・例(対象ドラマ:「下町ロケット」)

(a) タイトル・内容のインパクトがある(作業員 A・作業員 B で共通)

タイトル	ウェブページの種類	「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の組合せ	タイトルのインパクトの根拠となる文言	内容のインパクトの根拠
【2015年・下町ロケット】 佃製作所のロケ地(社屋)は大田区「桂川精螺製作所」	ブログ	ドラマ名 ロケ地	ロケ地	「ロケ地」情報を含むため
『下町ロケット』視聴率苦戦、原因は土屋太鳳か…多い女性アンチ、演技が反感買う	非ブログ	ドラマ名 主要な俳優 視聴率 評価表現	原因は土屋太鳳 演技が反感買う	「土屋太鳳は女性アンチが多い」という意外な事実のため

(b) タイトルのインパクトがある・内容のインパクトがない(作業員 A・作業員 B で共通)

タイトル	ウェブページの種類	「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の組合せ	タイトルのインパクトの根拠となる文言	内容のインパクトがない理由
『下町ロケット(2018)』2話感想 軽部がちよっと好きになってきた w	ブログ	ドラマ名 評価表現	軽部がちよっと好きになってきた w	他人の感想が羅列してあるだけで、まとまった意見が存在しないため
『下町ロケット』高視聴率も「下町ボブスレー」騒動のせいで白けてる? 「下町の技術力」という幻想?	非ブログ	ドラマ名 視聴率 評価表現	「下町ボブスレー」騒動のせいで白けてる?	「下町ボブスレー」が下町ロケットに関係ないため

(c) タイトルのインパクトがない・内容のインパクトがある(作業員 A, または, 作業員 B のみ)

タイトル	ウェブページの種類	「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の組合せ	タイトルのインパクトがない理由	内容のインパクトの根拠
なん J PRIDE: 落合福嗣くん、下町ロケットで俳優デビュー wtwtwtwtwtwtwuwtw	ブログ	ドラマ名	タイトル全体	プロ野球中日の元監督・落合博満氏(64)の長男であり、俳優デビューしたため
島津は仮面ライダーに部屋を守られている?	非ブログ	その他	タイトル全体	作品中に仮面ライダーシリーズのキャラクター名や俳優が多数存在しているのが、意外であるため

(d) タイトル・内容のインパクトがない(作業員 A・作業員 B で共通)

タイトル	ウェブページの種類	「ドラマ名・主要な俳優・ロケ地・視聴率・評価表現」の組合せ	タイトルのインパクトがない理由	内容のインパクトがない理由
【見逃し配信】『下町ロケット』1話動画ネタバレ感想と無料視聴する方法	ブログ	ドラマ名	見逃し配信	見逃し配信サービスを紹介するサイトであり、中身がないため
日曜劇場「下町ロケット」ヤタガラス4色ボールペン 通販	非ブログ	ドラマ名	通販	ボールペンを販売するECサイトのであり、中身がないため

り、本論文のタスクとは異なる。

6 おわりに

文献[6]の結果をふまえて、本論文では、文献[6]で訓練・評価対象としたウェブページ群の内容のインパクトの有無について分析を行った結果を述べるとともに、インパクト有無の自動判定手法を提案し評価を行った結果について述べた。具体的には、本論文では、ドラマ名をクエリとしたサジェストによって収集したブログ・非ブログのウェブページを対象として、ウェブページタイトルのインパクトの有無を判定するタスクを設定した。そして、このタスクに対して、事前訓練済みBERT[3]の

分散表現を用いた双方向LSTMを適用し、ウェブページタイトルのインパクト判定が一定の性能で実現できることを示した。

文 献

- [1] S. Chi, H. Luyao, and Q. Xipeng. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. In *Proc. NAACL-HLT*, pp. 380–385, 2019.
- [2] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, Vol. 20, No. 1, pp. 37–46, 1960.
- [3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proc. NAACL-HLT*, pp. 4171–4186, 2019.
- [4] X. Hu, L. Bing, S. Lei, and Y. Philip. BERT post-training for review reading comprehension and aspect-based sentiment analysis. In *Proc. NAACL-HLT*, pp. 2324–2335, 2019.

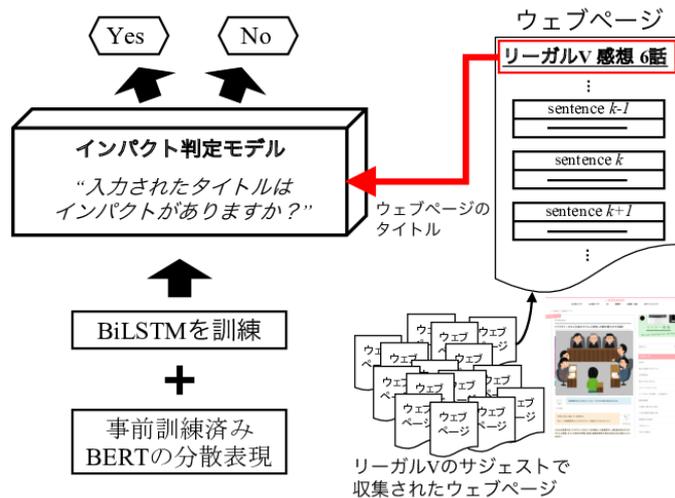
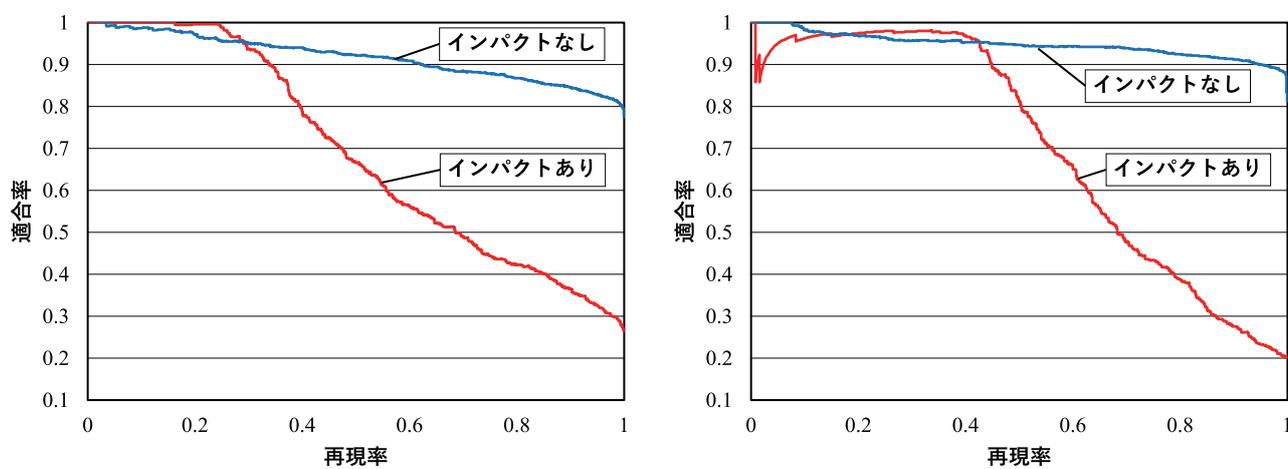


図4 インパクト判定モデル



(a) 作業者 A

(b) 作業者 B

図5 ウェブページタイトルのインパクト判定の評価結果

- [5] H. Jack and L. Lillian. Something's brewing! early prediction of controversy-causing posts from discussion features. In *Proc. NAACL-HLT*, pp. 1648–1659, 2019.
- [6] K. Kawaguchi, R. Kubo, T. Fujita, T. Maeda, T. Utsuro, A. Kobayashi, H. Nishizaki, and Y. Kawada. BERT based Web mining of concerns and reviews for TV drama audience. In *Proc. WI Workshops*, pp. 64–71, 2019.
- [7] C. Sihao, K. Daniel, Y. Wenpeng, C. Chris, and R. Dan. Seeing things from a different angle: discovering diverse perspectives about claims. In *Proc. NAACL-HLT*, pp. 542–557, 2019.
- [8] W. Xiaozhi, H. Xu, L. Zhiyuan, S. Maosong, and L. Peng. Adversarial training for weakly supervised event detection. In *Proc. NAACL-HLT*, pp. 998–1008, 2019.