

政治家のYouTubeを対象とした話者識別の定量的評価の試み

渡邊隆誠・木村泰知(小樽商科大学)・森浩太(株式会社JDSC)・加藤賢(VETA株式会社)・御器谷裕樹(慶應義塾大学)・吉田光男(筑波大学)・粕谷祐子(慶應義塾大学)

背景・目的

政治家による情報発信の手段として「動画」が定着

動画での発言内容について知りたいときに、動画データのままで、各政治家の言説が簡単には分からない

➡ 文字起こし、話者識別を自動化したい

目的

YouTubeの動画から政治家の発言を自動的に抽出できるかどうかという検証を行う

- WhisperXによる、文字起こしおよび話者識別の可否を検討
- 政治に関する動画分析における既存の話者識別手法の適用可能性を明らかにする

実験の流れ

(小泉-3)

WhisperXによる話者識別

[話者2]:賃金上昇に向けた政策を総動員していきたいと思います

[話者1]:それでは3列目の一番最初に手を挙げていただいた方よろしく申し上げますマイクお願いします
白タイムズの尾形ですおはようございますこんにちは

[話者3]:小泉さん、最初にですね、国民の不安のことに耳を傾ける

← 識別ミスした箇所

YouTube動画



混同行列	WhisperXによる話者識別			
	話者 1	話者 2	話者 3	
正解	話者 1	927	168	64
	話者 2	88	13,037	253
	話者 3	19	299	4,340

この動画(小泉-3)における各指標の値

マクロ再現率	マクロ適合率	マクロF値
0.902	0.9313	0.9157

正解

[話者2]:賃金上昇に向けた政策を総動員していきたいと思います

[話者1]:それでは3列目の一番最初に手を挙げていただいた方よろしく申し上げますマイクお願いします

[話者3]:**白タイムズの尾形ですおはようございます**

[話者2]:**こんにちは**

[話者3]:小泉さん、最初にですね、国民の不安のことに耳を傾ける

← 正解の話者識別

実験対象・結果

政党	動画ID	動画長	話者数	カテゴリ	マクロF値
共産	田村-1	1:55	2	演説	0.8682
共産	田村-2	0:59	6	ショート	0.1642
共産	田村-3	0:53	3	ショート	0.3649
みらい	安野-1	0:42	3	ショート	1
みらい	安野-2	0:20	1	ショート	1
みらい	安野-3	0:57	2	ショート	0.4662
維新	前原-1	16:39	2	対談	0.8696
維新	前原-2	0:52	1	ショート	1
維新	前原-3	27:36	5	議会	0.8307
公明	岡本-1	0:45	1	演説	1
公明	岡本-2	6:08	1	演説	1
公明	岡本-3	6:09	2	議会	1
国民	玉木-1	29:00	2	対談	0.8832
国民	玉木-2	9:59	3	密着	0.5610
国民	玉木-3	8:42	1	解説	1
参政	神谷-1	10:50	4	議会	0.7446
参政	神谷-2	10:36	2	解説	0.8703
参政	神谷-3	-	-	-	-
自民	小泉-1	1:37	2	ショート	0.4344
自民	小泉-2	7:58	4	議会	0.7233
自民	小泉-3	65:27	3	記者会見	0.9157

これらの動画で発言文字数を基に
マクロ再現率
マクロ適合率
マクロF値
の三指標を算出

誤り分析

(田村-2)

ショート動画は話者識別の精度が低い

誤識別例)

いろんな声を聞かせてもらえればと思って「奨学金を返しながらずっと最低線に働いてというのは本当に苦しくて」「借金背負って社会に出なかんっていうのはもうなくしたいって思う子供」だから返しながら自分の子供の教育費どうしようって考えるんだよね「一緒にいるのがすごい重要視されてる」みんなと一緒に安心民主主義は意見が違って当たり前なんです

上記の例はWhisperXが一人の話者による連続した発言と判断したが、実際には四人の話者による発言であった

このようにショート動画では話者数の過小推定により精度が著しく低下する事例が複数件確認された

考察

話者識別は実用的な精度なのか

・短い相槌など両者が同時に発話する区間において、主要話者の発言として一括して識別される事例

➡ 誤認識された区間は意見表明などの重要な内容を含む箇所ではなく、発言内容分析への影響は限定的

・ショート動画では、全体としてマクロF値が低い傾向

➡ ショート動画特有の高速なカット割りなどにより、話者クラストの分離が困難になり、誤識別、話者数の過少推定

「超短尺かつ多人数」の動画が適用範囲外となる可能性

結論

- 一部形式の動画では実用的な識別精度
- 政治家を直接は同定できない
- 今後は得られたデータを用いて発言の一貫性を検証