

MSc Project 2021

Title: Estimating personality in communication
Name: Yuichi Midorikawa

weekX

This indicates when it was done

X-X

This corresponds to the mindmap number

1. Summary of actions agreed during last meeting

1-1. I shared the time plan to manage my project.

Time Plan		Oct				Nov					Dec	My Progress	
Step	Task	4	11	18	25	1	8	15	22	29	6	Status	week2 Question (Yes/No)
		week1	week2	week3	week4	week5	week6	week7	week8	week9	week10		
1	capture a dataset that contains people talking and the text of what they say.	1					buffer				buffer	Finished	
	1-1. Find a dataset to use	1-1										Finished	
2	using the conversation text, do sentiment analysis (A)	2										Running	
	2-1. Find a model to use	2-1										Running	Yes
	2-2. Using the model and its data set, perform sentiment analysis		2-2									-	
3	from the videos, extract people and body pose (B)	3										Running	
	3-1. Find a model to use	3-1										Running	
	3-2. Using the model and its data set, extract body pose		3-2									-	
4	from the head, extract facial feature points (C)	4										Running	
	4-1. Find a model to use	4-1										Running	
	4-2. Using the model and its data set, extract facial points		4-2									-	
5	Then train a model to predict (A) from (B)+(C)			5								-	
	5-1. Predict (A) from (B) + (C)			5-1								-	
6	evaluate and analyse the results.				6							-	
	6-1. Decide a evaluation metrics				6-1							-	
	6-2. evaluate and analyse the results					6-2						-	
7	Write a paper							7				-	1-1

Research Steps

1. capture a dataset that contains people talking and the text of what they say.
2. using the conversation text, do sentiment analysis (A)
3. from the videos extract people and body pose (B)
4. from the head, extract facial feature points (C)
5. Then train a model to predict (A) from (B)+(C)
6. evaluate and analyse the results.

2. Summary of work done & results this week

2-1. I updated my research steps

2-2. I have selected a dataset to use (Step 1-1)

2-3. I extracted facial features using Openface from Youtube video.(Step 4-1)

Research Steps (Updated on week 2)

week2

2-1

Data Preparation

1-1. capture a dataset that contains people talking and the text of what they say.



Youtube



1-2. Watch the video and manually annotate each subtitle/frame with a positive/negative. (A') (Use as training data)



①

Positive/
Negative

Implementation

2. using the conversation text,
do sentiment analysis (A)



②

Positive/
Negative

3. from the videos, extract
people and body pose (B)



4. from the head, extract facial
feature points (C)



5. Then train a model to predict
(A) or (A') from (B)+(C)



③

Positive/
Negative

Evaluation

6. evaluate and analyse the
results.

①

Positive/
Negative

②

Positive/
Negative

③

Positive/
Negative

2-2. I have selected a dataset to use (Step 1-1)

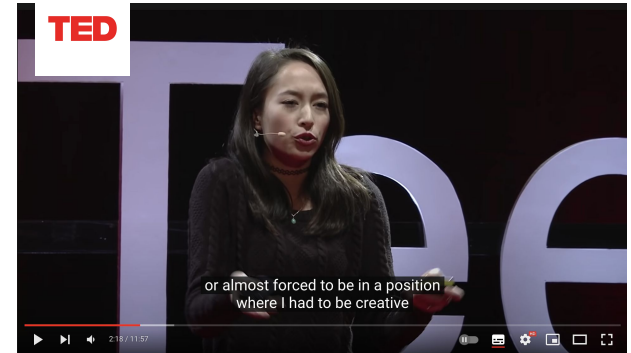
I got the TED video from Youtube to use in this project.

Title: [Why I Don't Use A Smart Phone | Ann Makosinski | TEDxTeen](#)

→ I would like to find some other videos with different conditions.

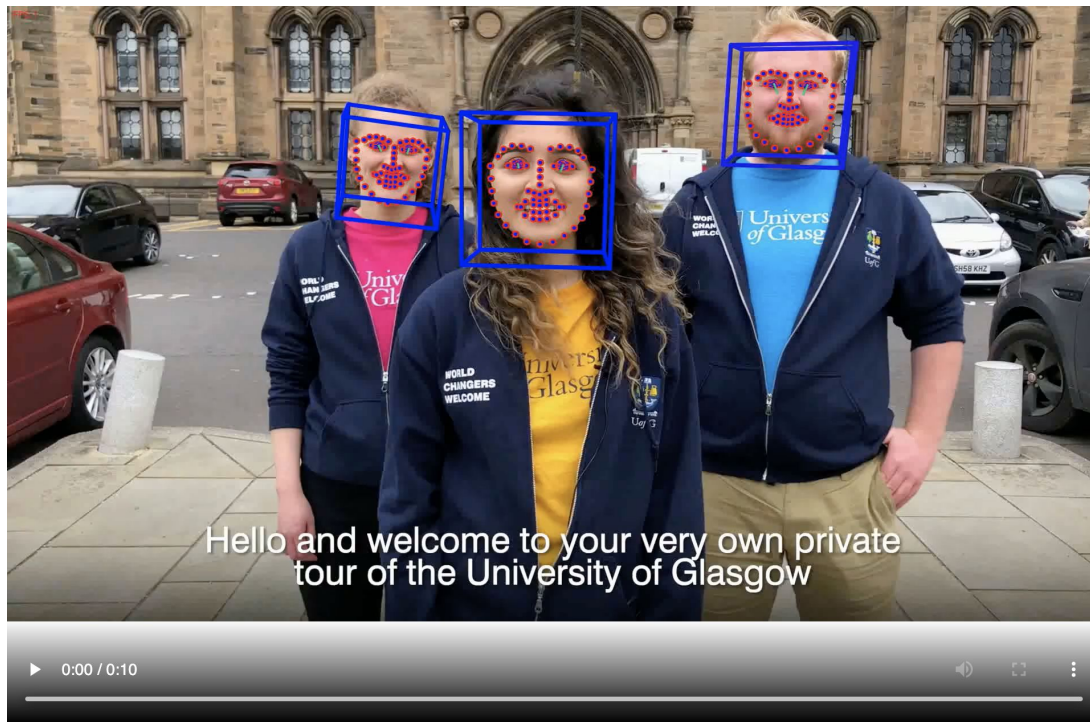
What is the ideal video for Youtube in this project?

- One or two people are talking.
- Facial expressions can be detected.
- Body posture can be detected.
- Emotional expression is as much as possible.
- English subtitles are available.



2-3. I extracted facial features using Openface from Youtube video.(Step 4-1,

Sample demo video: University of Glasgow campus tour



3. Questions to be discussed during the meeting

3-1. How many people do I need at least to manually annotate sentiment on a Youtube video? (depends on the video? 3 people?)

4. Proposed objectives for next week

4-1. I will explore technologies for (1)sentiment analysis from text, extracting (2)body posture and (3)facial features, and implement them. (step 2-2, 3-2, 4-2)

4-2. I will plan specific output/objective variables. (Positive/Negative/Neutral ?)

no need 4-3. According to the output, the Youtube video will be annotated manually.

<http://kahlan.eps.surrey.ac.uk/savee/>

<https://ethz.ch/content/dam/ethz/special-interest/baug/igp/photogrammetry-remote-sensing-dam/documents/pdf/schindler08nn.pdf>

4-1. I will explore technologies for (1)sentiment analysis from text, extracting (2)body posture and (3)facial features, and implement them. (step 2-2, 3-2, 4-2)

week2

Data Preparation

1-1. capture a dataset that contains people talking and the text of what they say.



Youtube



e.g. VIBE, etc

1-2. Watch the video and manually annotate each subtitle/frame with a positive/negative. (A') (Use as training data)



e.g. OpenFace, etc

Negative

Implementation

2. using the conversation text, do sentiment analysis (A)



②

Positive/
Negative

3. from the videos, extract people and body pose (B)



5. Then train a model to predict (A) or (A') from (B)+(C)



4. from the head, extract facial feature points (C)



③

Positive/
Negative

Evaluation

6. evaluate and analyse the results.

①

Positive/
Negative

②

Positive/
Negative

③

Positive/
Negative

5. Articles read this week

5-1. Estimating the Intensity of Facial Expressions Accompanying Feedback Responses in Multiparty Video-Mediated Communication

<https://dl.acm.org/doi/abs/10.1145/3382507.3418878>

5-2. Predicting Influential Statements in Group Discussions using Speech and Head Motion Information

<https://dl.acm.org/doi/10.1145/2663204.2663248>

End

Youtubeの加工方法の方針

1. Youtubeをダウンロード

センチメント分析の出力結果を定義する

Research Steps (Updated on week 2)

week2

2-1

Data Preparation

1-1. capture a dataset that contains people talking and the text of what they say.



Youtube



1-2. Watch the video and manually annotate each subtitle/frame with a positive/negative. (A') (Use as training data)



①
Positive/
Negative

Implementation

2. using the conversation text,
do sentiment analysis (A)



②
Positive/
Negative

3. from the videos, extract
people and body pose (B)



4. from the head, extract facial
feature points (C)



5. Then train a model to predict
(A) or (A') from (B)+(C)



③
Positive/
Negative

Evaluation

6. evaluate and analyse the
results.

①
Positive/
Negative

②
Positive/
Negative

③
Positive/
Negative

Implementation -
using the conversation text, do sentiment analysis (A)

I could conduct sentiment analysis by using VADER.

```
┌→      neg    neu    pos  compound
I am happy. 0.000 0.213 0.787  0.5719
I am sad.   0.756 0.244 0.000 -0.4767
I am angry. 0.767 0.233 0.000 -0.5106
```

Research Title and Abstract

Title: Estimation of personality/sentiment in communication

Abstract:

1. I obtain (1) text data (content of conversation) and (2) hand gesture (or face) from videos of people talking.
2. Using the data in (1), perform (3) sentiment extraction (e.g., Positive, Negative, Neutral).
3. Using the data from (2) and (3), evaluate and analysis the both sentiment in communication.

VIBE: Video Inference for Human Body Pose and Shape Estimation [CVPR-2020]

<https://github.com/mkocabas/VIBE>

Emotion Classification in Short Messages

<https://github.com/lukasgarbas/nlp-text-emotion>

Data Set

こういったデータセットなのか？（概要を使って説明）

具体的にはこういったデータがあるのか？（テーブルを使って説明）

このデータを使って何をするのか？（＝テーマを改めて説明）

それをこういったスケジュールで進めるのか？

今はこういった作業をしているのか？

【Python】日本語による感情分析をTransformersで行う

HuggingfaceのTransformersをインストールする

<https://self-development.info/huggingface%E3%81%AEtransformers%E3%82%92%E3%82%A4%E3%83%B3%E3%82%B9%E3%83%88%E3%83%BC%E3%83%AB%E3%81%99%E3%82%8B/>

【Python】Mecabのラッパであるfugashiのインストール

<https://self-development.info/%E3%80%90python%E3%80%91mecab%E3%81%AE%E3%83%A9%E3%83%83%E3%83%91%E3%83%BC%E3%81%A7%E3%81%82%E3%82%8BFugashi%E3%81%AE%E3%82%A4%E3%83%B3%E3%82%B9%E3%83%88%E3%83%BC%E3%83%AB/>

【Python】日本語による感情分析をTransformersで行う

<https://self-development.info/%E3%80%90python%E3%80%91%E6%97%A5%E6%9C%AC%E8%AA%9E%E3%81%AB%E3%82%88%E3%82%8B%E6%84%9F%E6%83%85%E5%88%86%E6%9E%90%E3%82%92transformers%E3%81%A7%E8%A1%8C%E3%81%86/>

オプション

- ・日本語と英語のテキストを使って、センチメント分析を行う

→これってYoutubeの字幕の精度の研究になっちゃう？

（共通のセンチメント分析のライブラリを使用できるのか？もし違うなら前提が違くなるから、あまり意味がない？）

データセット

成蹊大学のデータセットはなぜ利用しづらいのか

(1)動画がないから人手によるアノテーションがしづらい

→教師データを新しく作りづらい

(2)既存のデータを教師データとして扱うことは可能である。

(3)対話テキストが少し日本語としておかしい箇所がある。

代替案

アイデア

- (1) fake news (テキスト)
- (2) サマリー生成 超長い文章→要約(テキスト)
- (3) 文章生成(テキスト)
- (4) 画像生成(画像)
- (5) 画像+テキスト: boketeの英語版
- (6) 芸術系 この絵に対して、どう思う？(アノテーションが必要)
- (7) コミュニケーション系+画像

代替案

Real Life Violence Situations Dataset

<https://www.kaggle.com/mohamedmustafa/real-life-violence-situations-dataset>

icon

<https://icooon-mono.com/>