

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

Time Plan		Oct				Nov					Dec	My Progress	
Step	Task	4	11	18	25	1	8	15	22	29	6	Status	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										Finished	
	2-1. Find a model to use	2-1										Finished	
	2-2. Using the model and its data set, perform sentiment analysis		2-2									Finished	
3	from the videos, extract people and body pose (B)	3										Finished	
	3-1. Find a model to use	3-1										Finished	
	3-2. Using the model and its data set, extract body pose		3-2									Running	
4	from the head, extract facial feature points (C)	4										Finished	
	4-1. Find a model to use	4-1										Finished	
	4-2. Using the model and its data set, extract facial points		4-2									Finished	
5	Then train a model to predict (A) from (B)+(C)			5								Running	
	5-1. Predict (A) from (B) + (C)			5-1								Running	
6	evaluate and analyse the results.				6							Running	
	6-1. Decide a evaluation metrics				6-1							Running	
	6-2. evaluate and analyse the results					6-2						-	
7	Write a paper							7				Running	

Research Steps (Updated on week 4)

week4

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)



Positive/
Negative

Implementation (Using a pre-trained model)

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



①

Positive/
Negative

Use as training
data

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)

②

Positive/
Negative

Evaluation

6. evaluate and analyse the
results.

①

Positive/
Negative

②

Positive/
Negative

Evaluation metrics:
Accuracy,
Precision,
Recall, and
F1-score...

1. Summary of actions agreed during last meeting

1-1. Decided on Emotions Category for this project

- Plan A. Ekman (anger, disgust, fear, joy, sadness, surprise + neutral)
- Plan B. Hierarchical Grouping (positive, negative, ambiguous + neutral)

1-2. Decided on Text Emotion Classification Model (BERT based model)

2-1. Decided on Emotions Category for this project

According to the following paper, the emotion categories **Hierarchical Grouping** and **Ekman** were high F1-score, so for now I will use **these emotion categories** in this project.

I would like to determine Hierarchical Grouping or Ekman according to the emotional categories that **the facial expressions and body postures can output**.

Emotion category name	Number of emotions	F1-score	
Original GoEmotions	27 emotions + neutral	0.46	
Hierarchical Grouping	positive, negative, ambiguous + neutral	0.69	2
Ekman	anger, disgust, fear, joy, sadness, surprise + neutral	0.64	1

[1] GoEmotions: A Dataset of Fine-Grained Emotions

<https://arxiv.org/abs/2005.00547>

[2] P. Ekman, Universal facial expressions of emotion 151–158.

2-2. Decided on Text Emotion Classification Model

The top performing models in the EmotionX Challenge (Hsu and Ku, 2018) all used the pre-trained BERT model.

→ Therefore, I plan to use the **BERT** model in my project as well. I will use an existing, pre-trained model called Go-emotion.

Model	Embeddings	Macro-F1
SVM	-	0.55
Random Forest	-	0.49
BiLSTM+CNN+ Self-Attention	GoobleEmb	0.62
	GloVE	0.62
	FastText	0.63

Table 1: Comparison of BiLSTM+CNN+Self-Attention models for ISEAR

Model	Macro-F1
BERT	0.702
RoBERTa	0.742
DistilBERT	0.693
XLNet	0.731

Table 3: Comparison of Transformers models for ISEAR

- [1] Chao-Chun Hsu and Lun-Wei Ku. 2018. SocialNLP 2018 EmotionX Challenge Overview: Recognizing Emotions in Dialogues
https://www.researchgate.net/publication/328137931_SocialNLP_2018_EmotionX_Challenge_Overview_Recognizing_Emotions_in_Dialogues
- [2] Exploring Transformers in Emotion Recognition: a comparison of BERT, DistilBERT, RoBERTa, XLNet and ELECTRA
<https://arxiv.org/abs/2104.02041>

2. Summary of work done & results this week

2-1. Perform sentiment analysis on each sentence of youtube videos

2-2. Extract facial feature points from the head

2-3. Organize the analysis procedure

2-1. Perform sentiment analysis on each sentence of a youtube video

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



Basic information:

Viewing time: 11:57

The number of sentences: 209

Emotions Category	Number of emotion	(%)
neutral	129	62%
joy	52	25%
surprise	17	8%
anger	6	3%
sadness	4	2%
fear	1	0%
disgust	0	0%
Total	209	-

2-1. Perform sentiment analysis on each sentence of a youtube video

Video Title: [How to speak so that people want to listen | Julian Treasure](#)



Basic information:

Viewing time: 9:58

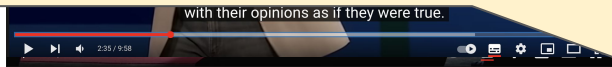
The number of sentences: 228

Emotions Category	Number of emotion	(%)
neutral	145	64%
joy	41	18%
anger	17	8%
surprise	16	7%
sadness	7	3%
disgust	2	1%
fear	0	0%
Total	228	-

2-1. Perform sentiment analysis on each sentence of a youtube video

Next Actions

- The emotion categories were uneven: there were few emotions other than 'neutral'.
- I'll try to add different videos to increase the number of other emotions.



Basic information:

Viewing time: 9:58

The number of sentences: 228

Emotions Category	Number of emotion	(%)
neutral	145	64.0%
joy	41	18.0%
anger	17	8.0%
surprise	16	7.0%
sadness	7	3%
disgust	2	1%
fear	0	0%
Total	228	-

2-1. Perform sentiment analysis on each sentence of a youtube video

	start	end	diff	script	emotion	score
100	05:57:00	06:03:00	00:06:00	But I never felt like, "Oh my God, I'm missing out by not having a phone."	sadness	0.969291
101	06:03:00	06:06:00	00:03:00	And, as some teenagers here may know, it's called "FOMO,"	neutral	0.999004
102	06:06:00	06:08:00	00:02:00	which is: Fear Of Missing Out.	neutral	0.907724
103	06:08:00	06:12:00	00:04:00	I never had that because I was so content with what I was given	joy	0.841819
104	06:12:00	06:14:00	00:02:00	and how much more I had to pursue.	joy	0.99628
105	06:15:00	06:18:00	00:03:00	So, what did I do in my elementary school?	surprise	0.992093
106	06:18:00	06:20:00	00:02:00	Well, when I was in elementary school,	neutral	0.999003
107	06:20:00	06:24:00	00:04:00	I was definitely not the coolest kid."	anger	0.979123
108	06:24:00	06:27:00	00:03:00	I was not the person who wanted to hang out with them."	neutral	0.982704
109	06:27:00	06:30:00	00:03:00	Because first of all, in middle school and high school	neutral	0.999097
110	06:30:00	06:32:00	00:02:00	you are really judged a lot,	neutral	0.932911
111	06:32:00	06:36:00	00:04:00	and I was very unconfident, at first, of how I appeared.	surprise	0.869203
112	06:36:00	06:39:00	00:03:00	I had short hair, glasses, braces. I dressed in guys' clothes.	neutral	0.998868
113	06:39:00	06:41:00	00:02:00	I didn't have the coolest stuff.	anger	0.995947
114	06:41:00	06:45:00	00:04:00	People would come up to me and be like, "Oh, what a handsome boy you are!"	joy	0.998924
115	06:45:00	06:48:00	00:03:00	and I'd be like, "Thanks!" and just walk away.	joy	0.997454

(Reference) Youtube videos to use for this project

#N	Title	Number of sentences	neutral	joy	surprise	anger	sadness	fear	disgust
	TED youtube videos								
1	Why I Don't Use A Smart Phone Ann Makosinski TEDxTeen	209	129 (62%)	52 (25%)	17 (8%)	6 (3%)	4 (2%)	1 (0%)	0 (0%)
2	How to speak so that people want to listen Julian Treasure	228	145 (64%)	41 (18%)	16 (7%)	17 (8%)	7 (3%)	0 (0%)	2 (1%)
3	How to spot a liar Pamela Meyer	405	288 (71%)	53 (13%)	18 (4%)	32 (8%)	12 (3%)	2 (0%)	0 (0%)
4	Robert Waldinger: What makes a good life? Lessons from the longest study on happiness TED	212	148 (70%)	32 (15%)	18 (8%)	1 (0%)	12 (6%)	1 (0%)	0 (0%)
5	How to make stress your friend Kelly McGonigal	284	178 (62%)	71 (25%)	20 (7%)	6 (2%)	5 (2%)	4 (1%)	0 (0%)
6	How to Get Your Brain to Focus Chris Bailey TEDxManchester	306	227 (74%)	41 (13%)	18 (6%)	11 (4%)	7 (2%)	0 (0%)	2 (1%)
7									

This table shows the results of the sentiment analysis performed on each sentence of the Youtube video.
 In every video, 'neutral' accounted for more than 60%.
 → I'm going to try to find some videos other than TED.

2-2. from the head, extract facial feature points

I will use OpenFace which has close to human accuracy and allows me to get the facial feature 128-dimensional representation.

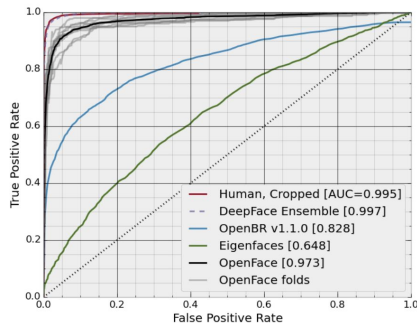


Figure 8: ROC curve on the LFW benchmark with area under the curve (AUC) values.



2-3. Analysis procedure (Updated on week4)

week4

step	Action item	Manual/Automatic	
1	Find a Youtube video	Manual	
2	Download the video	Automatic	01_youtube_downloader.ipynb
3	Get subtitles for the video and paste them into Excel	Manual	
4	Organise start times and sentences on one line in Excel	Automatic	02_organized_excel_file.ipynb
5	Find the end time and the difference (end time - start time)	Automatic	03_adding_time_information.ipynb
6	Looking for emotion in every sentence	Automatic	04_emotion_extraction_from_subtitles.ipynb
7	Cut out videos (mp4, avi) using playing time and differences	Automatic	
8	Extract facial feature points from the head using OpenFace	Automatic	
9	Extract facial feature points from the head using OpenPose	Automatic	
10	...		
11	...		
12	...		

3. Questions to be discussed during the meeting

3-1. What does the professor expect from students in this project in particular?

Is it to improve the accuracy of the model? Or that I am able to apply my comprehensive knowledge of data science, from acquiring the data to processing and analysing it?

4. Proposed objectives for next week

- 4-1. I will explore and implement techniques to extract features from (1) body posture. (Step 3-2)
- 4-2. Create a model to predict emotions and output the accuracy
- 4-3. improve the accuracy (Repeat this process!)

5. Articles read this week

5-1. OpenFace: A general-purpose face recognition library with mobile applications

<https://elijah.cs.cmu.edu/DOCS/CMU-CS-16-118.pdf>

5-2. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields

<https://arxiv.org/abs/1812.08008>

End