Inference of Student Needs in an Online Learning Environment Based on Facial Expression

Affective Engineering and Computer Arts Laboratory

Lee Jui Chi

Problem

- A teacher's teaching behavior has a significant impact on the motivation and involvement of students [1]
- Due to the limited interaction between students and teachers, teachers often struggle to keep track of their students' status in terms of the teaching response they need [2]
- Purpose: teachers will be able to know the exact student needs and adjust the teaching in real-time

[1] Skinner and Belmont, Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year, Journal of Educational Psychology
[2] Nailufar al. (2021) Analysis of Teacher Difficulties in Online Learning on Mathematics Subjects. Prisma Sains: Jurnal

Proposed Solution

- The objective of this research: to present a model for automatically inferring student needs based on facial expressions.
- The teacher can know how to respond to the student needs
- The teacher can know student learning status from the need
- The teachers can adjust their pace according to the needs to have an overall better educational experience

List of Student Needs

- Please teach faster [3]
- Please teach slower [3]
- Please wait a moment [3]
- Please skip this part [3]
- Please go back to the last part [3]
- Please let me ask a question [4]
- Please make the class more interesting [4]
- I don't need anything; please continue [4]
- Please explain more [5]
- Please let me take a break [5]

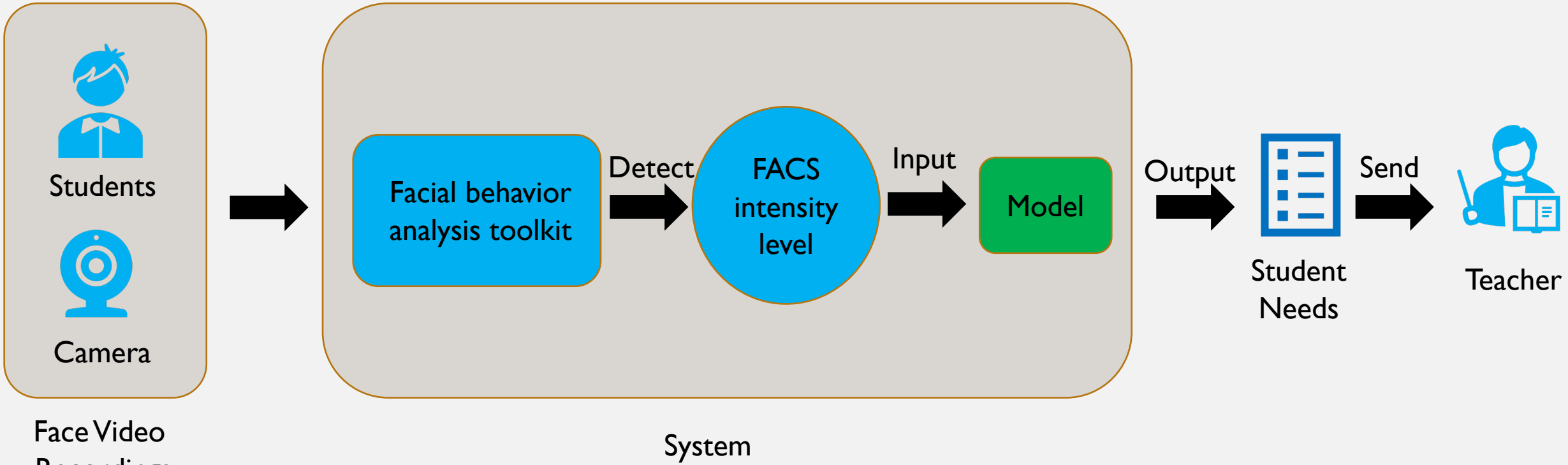
[3] Patel al. , Teach Me Fast: How to Optimize Online Lecture Video Speeding for Learning in Less Time?. In Proceedings of the Sixth International Symposium of Chinese CHI
[4] Apelojg (2018). Whats going on?! Needs and Emotions during Classes. The Felix-App: New Ways of Feedback and Evaluating Classes in Real-Time. *Conference on Society and Information Technologies*
[5] Nailufar (2021)

Facial Action Coding System

- The Facial Action Coding System (FACS) encodes a set of facial muscle movements was developed [6]
- The combinations of different facial actions and their intensity levels may reveal useful information for the inference model
- The intensity level (0.0 to 5.0) indicates how strongly the emotion is detected
- OpenFace recognizes 17 of 46 FACS Action Units (AU)

[6] Ekman, Friesen, Hager (2002). Facial Action Coding System:The Manual on CD ROM. Salt Lake City:A Human Face.

Proposed System Design



Student Needs Survey Experiment

Start to watch the video

2 mins

System pause the video

Report the needs

• • •

Click the play button to continue watching the video

2 mins

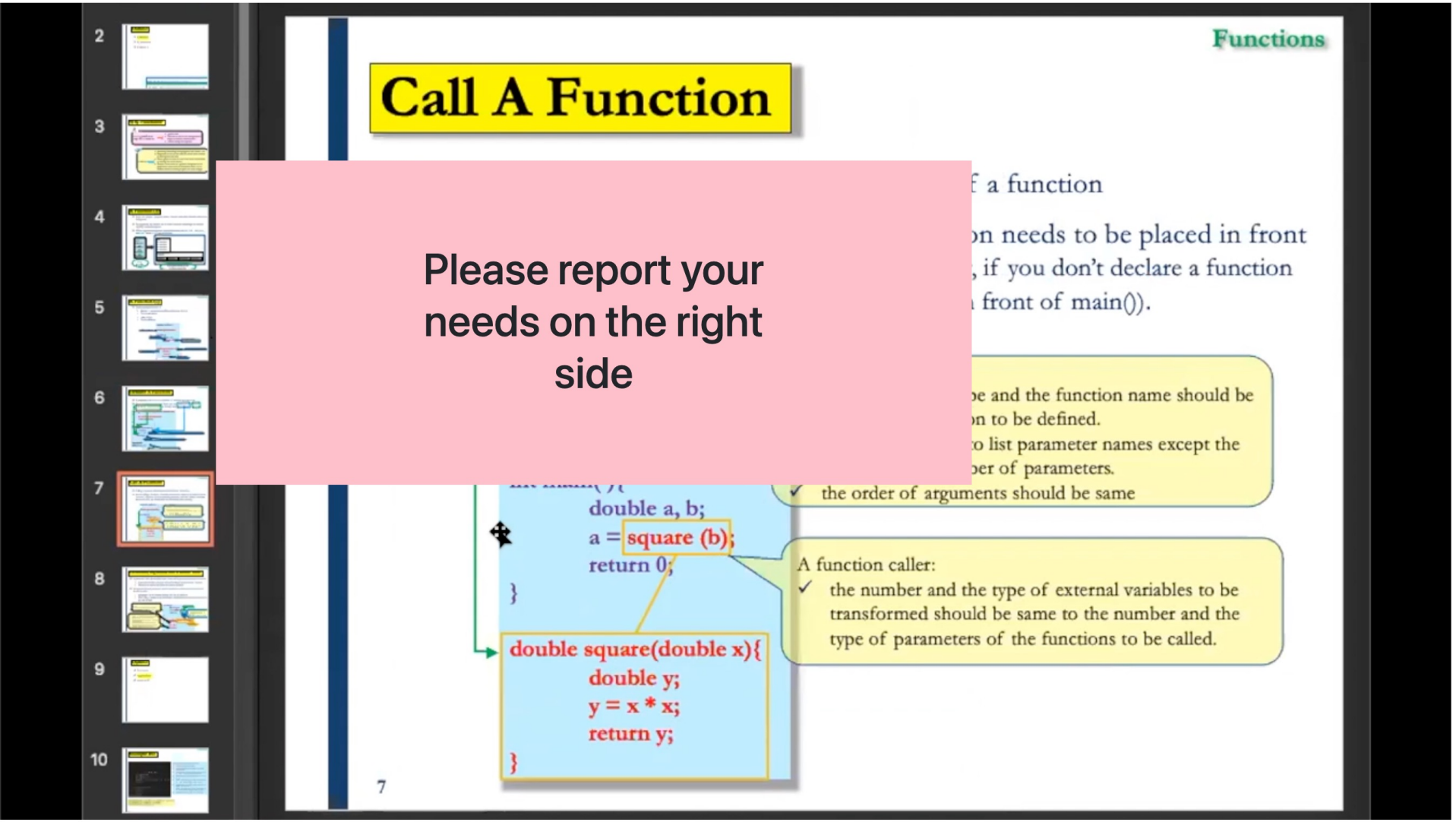
System pause the video

Report the needs

• • •

Repeat 4 times

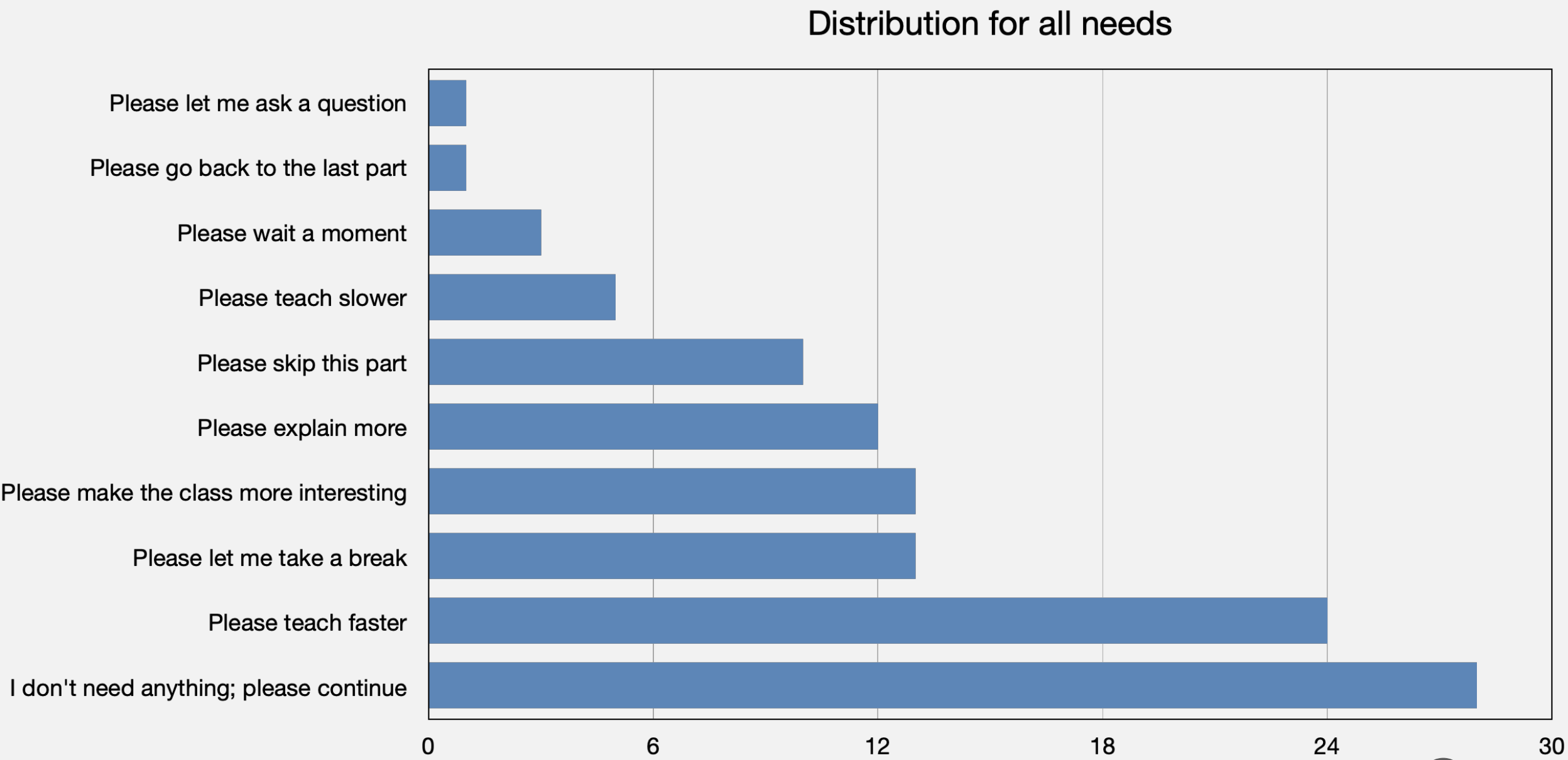
Play Video



Answer your needs at this point

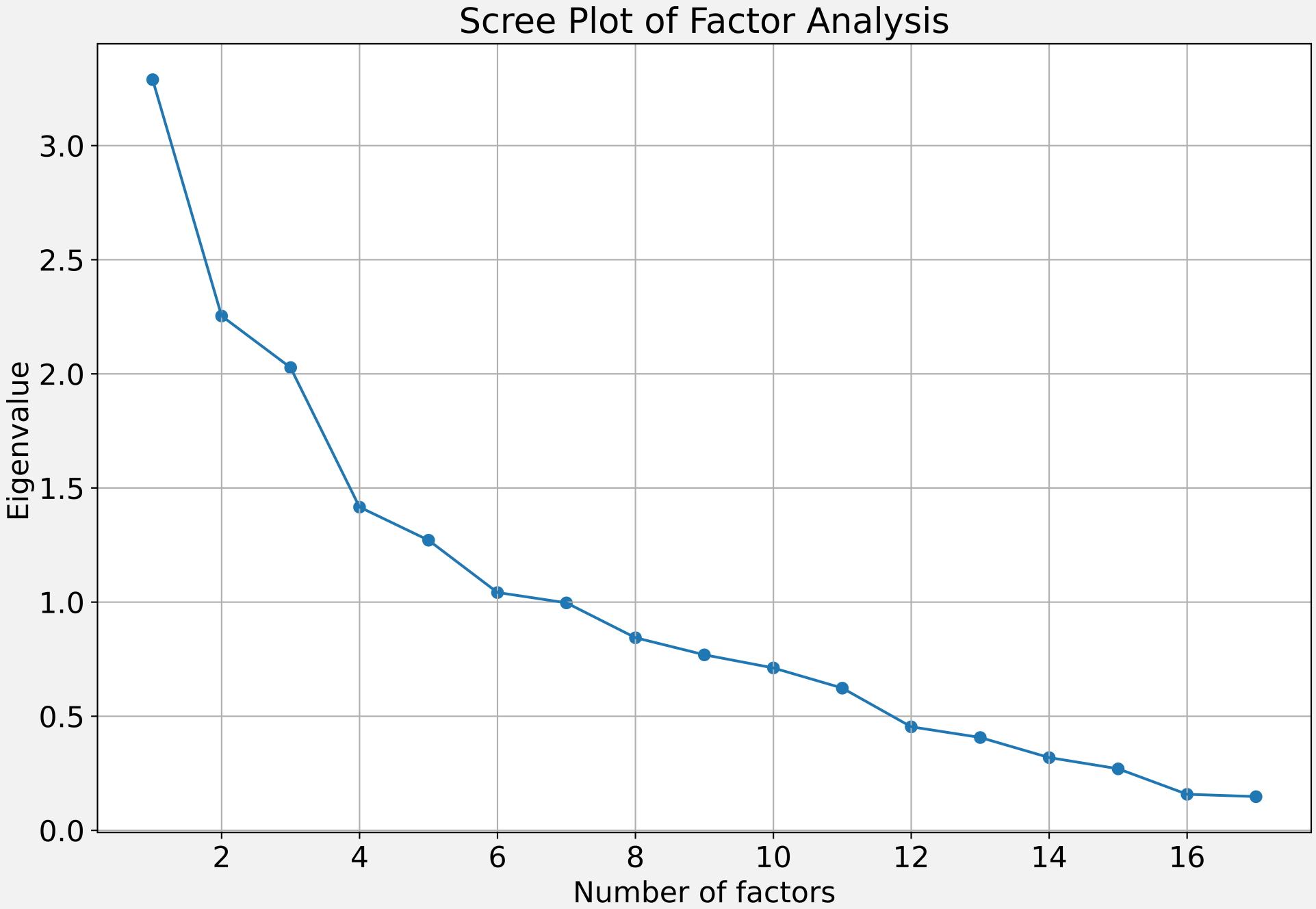
- [ ] I don't need anything; please continue
- [ ] Please teach faster
- [ ] Please teach slower
- [ ] Please wait a moment
- [ ] Please explain more
- [ ] Please let me take a break
- [ ] Please let me ask a question
- [ ] Please make the class more interesting
- [ ] Please skip this part
- [ ] Please go back to the last part

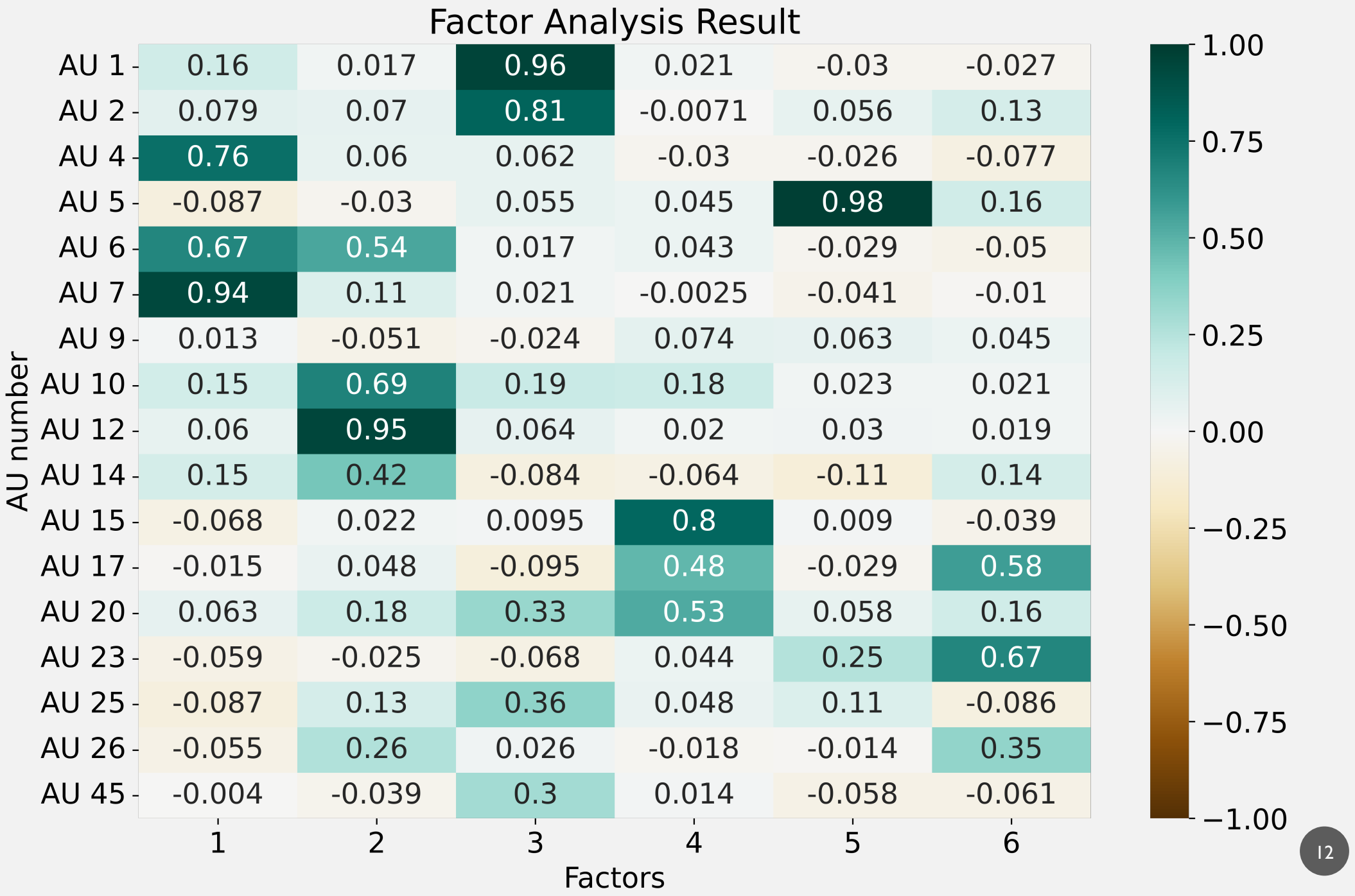
Submit


Distribution for all needs
Please let me ask a question
Please go back to the last part
Please wait a moment
Please teach slower
Please skip this part
Please explain more
Please make the class more interesting
Please let me take a break
Please teach faster
I don't need anything; please continue
0
6
12
18
24
30

Factor Analysis

- Factor Analysis on AUs was conducted to reduce the number of inputs and determine the most effective
- The result of the Kaiser-Meyer-Olkin (KMO) test revealed that it is meaningful to do the factor analysis
- Based on the Scree Plot and Kaiser Criterion result, the number of relevant factors that should be considered in the analysis was set to 6
- Varimax rotation to reduce the complexity


Scree Plot of Factor Analysis
Eigenvalue
Number of factors
0.0
0.5
1.0
1.5
2.0
2.5
3.0
2
4
6
8
10
12
14
16


Factor Analysis Result
AU number
AU 1 0.16 0.017 0.96 0.021 -0.03 -0.027
AU 2 0.079 0.07 0.81 -0.0071 0.056 0.13
AU 4 0.76 0.06 0.062 -0.03 -0.026 -0.077
AU 5 -0.087 -0.03 0.055 0.045 0.98 0.16
AU 6 0.67 0.54 0.017 0.043 -0.029 -0.05
AU 7 0.94 0.11 0.021 -0.0025 -0.041 -0.01
AU 9 0.013 -0.051 -0.024 0.074 0.063 0.045
AU 10 0.15 0.69 0.19 0.18 0.023 0.021
AU 12 0.06 0.95 0.064 0.02 0.03 0.019
AU 14 0.15 0.42 -0.084 -0.064 -0.11 0.14
AU 15 -0.068 0.022 0.0095 0.8 0.009 -0.039
AU 17 -0.015 0.048 -0.095 0.48 -0.029 0.58
AU 20 0.063 0.18 0.33 0.53 0.058 0.16
AU 23 -0.059 -0.025 -0.068 0.044 0.25 0.67
AU 25 -0.087 0.13 0.36 0.048 0.11 -0.086
AU 26 -0.055 0.26 0.026 -0.018 -0.014 0.35
AU 45 -0.004 -0.039 0.3 0.014 -0.058 -0.061
1 2 3 4 5 6
Factors
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Factor Analysis Result

* Using rules of thumb (0.55, good) for picking variables for factors [7]

- Factor 1: AU 4 (Brow Lowerer), AU 6 (Cheek Raiser), AU 7 (Lid Tightener)
- Factor 2: AU 10 (Upper Lip Raiser), AU 12 (Lip Corner Puller)
- Factor 3: AU 1 (Inner Brow Raiser), AU 2 (Outer Brow Raiser)
- Factor 4: AU 15 (Lip Corner Depressor)
- Factor 5: AU 5 (Upper Lid Raiser)
- Factor 6: AU 17 (Chin Raiser), AU 23 (Lip Tightner)

[7] Comrey, and Lee (2013). *A first course in factor analysis*. Psychology press.

Neural Network

- Double layer, 6 input nodes, and 10 outputs for the needs response
- 6 Inputs for the intensity levels of AUs by factors were normalized

$$x_F = \sum_i AU_i w_{Fi}$$

- AU_i is the normalized [0,1] AU intensity reported in the video frame for the ith AU in the model and w_{Fi} is a weight based on the factor loading of AU_i for factor F obtained by factor analysis

Output

- The average AU intensity level for video clips is used for testing
- One-tenth of the data were randomly chosen to be the test sets
- The outputs are floating values (from 0 to 1) for estimating student needs
- 0.5 is the criterion for determining whether it counts as showcasing the need, y_i is the network output

$$f(x) = \begin{cases} 1 & y_i \geq 0.5 \\ 0 & otherwise \end{cases}$$

- The training and test were repeated 50 times to obtain the average inference outcome that would serve as the result of this study.

Error Computation

- The difference between the reported needs and inferred needs is the error. y_i is infer needs, a_i is reported needs (0 or 1)

$$error = |y_i - a_i|$$

- If the error is bigger than 0.5, it is counted as a wrong inference result

$$Normalized\ error = \begin{cases} 1 & error > 0.5 \\ 0 & error \leq 0.5 \end{cases}$$

Result

Student needs	Average error
Please teach faster	0.383
Please explain more	0.222
Please make the class more interesting	0.229
Please let me take a break	0.275
Please skip this part	0.151
I don't need anything; please continue	0.389
Please teach slower	0.099
Please wait a moment	0.074
Please let me ask a question	0.02
Please go back to the last part	0.018

Table 1: Unweighted model inference average error

Average 1.35 wrong inference

Student needs	Average error
Please teach faster	0.377
Please explain more	0.212
Please make the class more interesting	0.281
Please let me take a break	0.277
Please skip this part	0.184
I don't need anything; please continue	0.4
Please teach slower	0.1
Please wait a moment	0.055
Please let me ask a question	0.023
Please go back to the last part	0.034

Table 2: Weighted model inference average error

Average 1.3475 wrong inference

Conclusions

- Developed a model to infer the student needs based on the facial expression
- The result demonstrated that it could correctly infer about 87% of the time
- Future work: improve the model's accuracy, increase the data, develop a user interface on the teacher's side