Improving Continuous Emotion Annotation in Video Platforms via Physiological Response Profiling

Author(s): **Swarnali Banik**, Sougata Sen, Snehanshu Saha, Surjya Ghosh

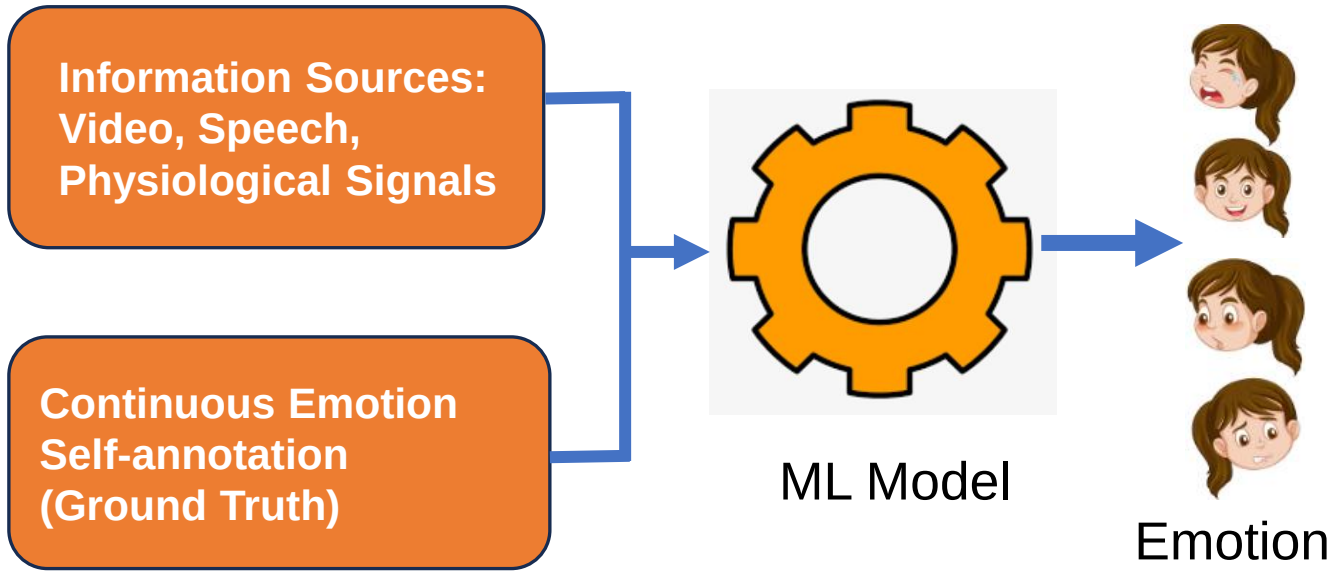
Computer Science and Information Systems, APPCAIR

BITS Pilani Goa, INDIA

12th International Conference on AFFECTIVE COMPUTING & INTELLIGENT INTERACTION (ACII 2024)
15th - 18th September 2024, Glasgow, UK

Motivation and Background

- Video-based applications
- Emotion Detection Using ML Models : A Multi-Modal Approach
  - Integrating Multiple Modalities
  - Emotion ground truth



Typical ML-based emotion detection model

Motivation and Background

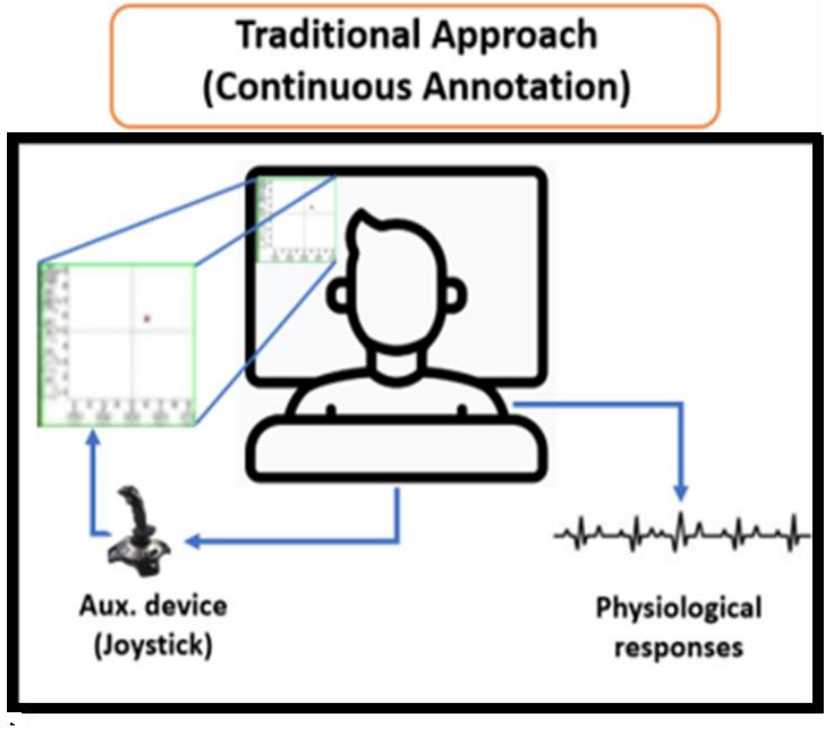
Emotion Self-Annotation

- **Post-interaction/ post stimuli**
  - Fails to capturing subtle intra-video nuances.

Motivation and Background

Emotion Self-Annotation

- **Continuous annotation**
  - Degrades viewing experience.
  - Increases cognitive load, leading to less accurate emotion labels.



Participant watching videos and annotate continuously

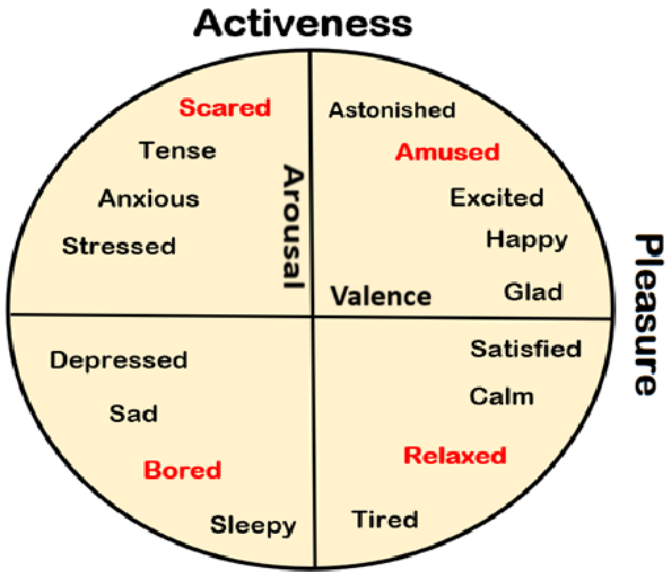
Research Question

Can we developed an intelligent annotation framework ?

- Capture ***opportune moments*** for emotion self-report collection
- Reducing the number of self-annotations (probes)
  - Decreases the cognitive workload

Dataset : CASE [1]

- 30 users (15 M, 15F)
- 2D plane - Joystick Input
- Collected emotion self-report annotation



Circumplex Model

Participant watching videos and annotate continuously

[1] - Sharma, K., Castellini, C., van den Broek, E.L. et al. A dataset of continuous affect annotations and physiological signals for emotion analysis. Sci Data 6, 196 (2019). https://doi.org/10.1038/s41597-019-0209-0

Dataset : CASE [1]

- 8 videos in randomized order
- Physiological sensors data collected
  - Electrocardiograph (ECG)
  - Blood Volume Pulse (BVP)
  - **Galvanic Skin Response (GSR)**
  - Respiration (RSP)
  - Skin Temperature (SKT)
  - Electromyography (EMG)

Video id	Emotion	Valence	Arousal	Duration (in sec.)
1	amusing	med/high	med/high	185
2	amusing	med/high	med/high	173
3	boring	low	low	119
4	boring	low	low	160
5	relaxing	med/high	low	145
6	relaxing	med/high	low	147
7	scary	low	high	197
8	scary	low	high	144

The 8 videos present in the CASE dataset

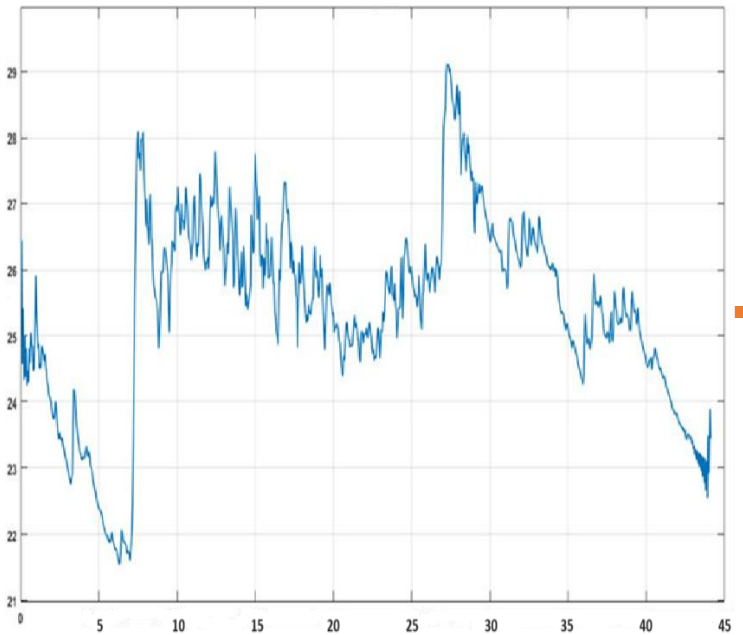
[1] - Sharma, K., Castellini, C., van den Broek, E.L. et al. A dataset of continuous affect annotations and physiological signals for emotion analysis. Sci Data 6, 196 (2019). https://doi.org/10.1038/s41597-019-0209-0

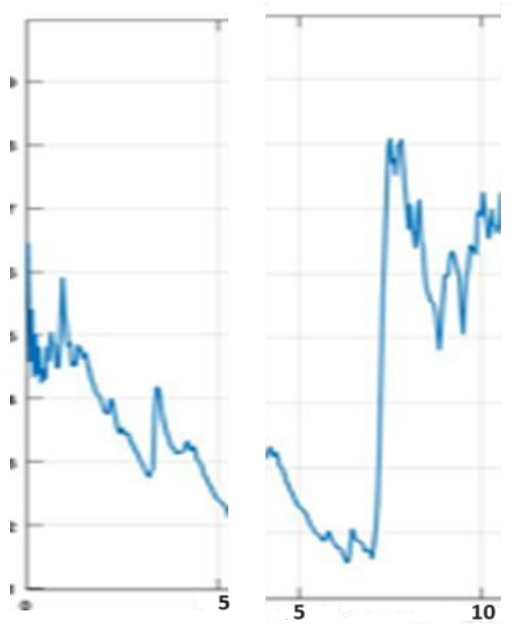
Labeling Opportune Moments

Physiological Responses

Segment signals into fixed size windows (5 second)

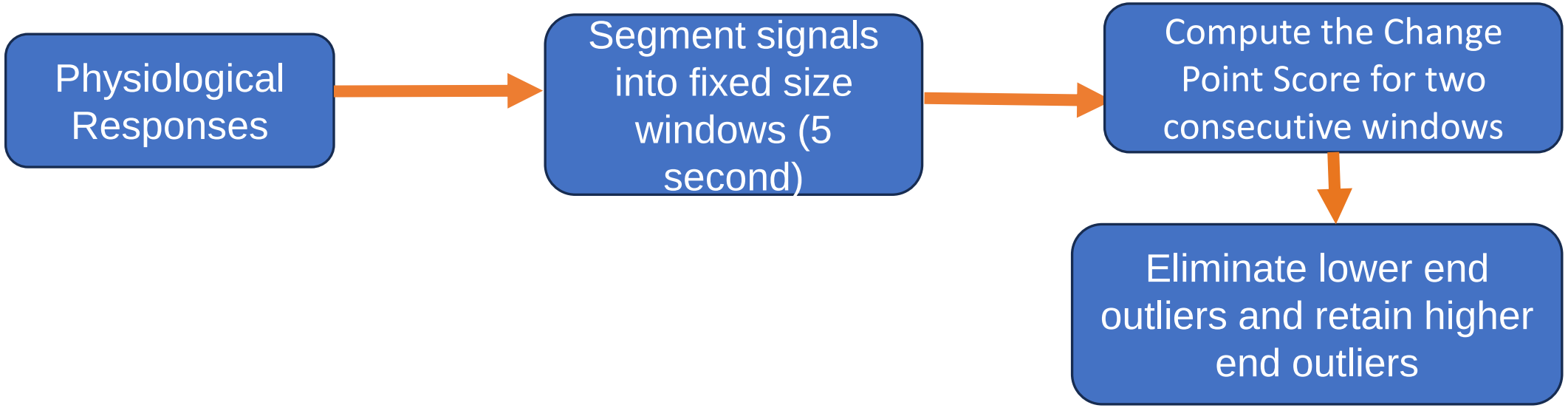
Compute the Change Point Score for two consecutive windows



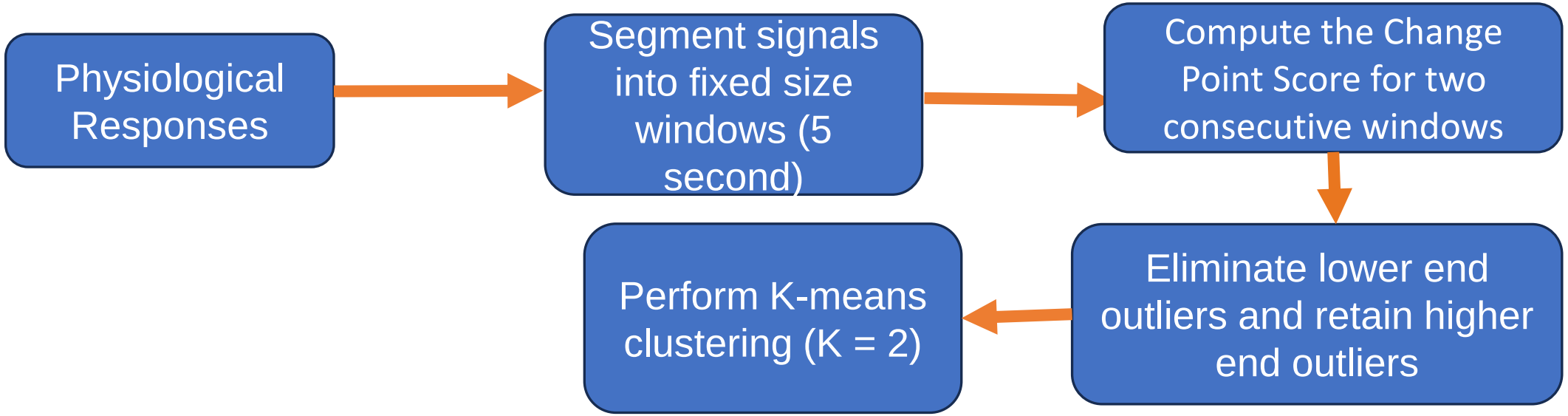


Change Point Score

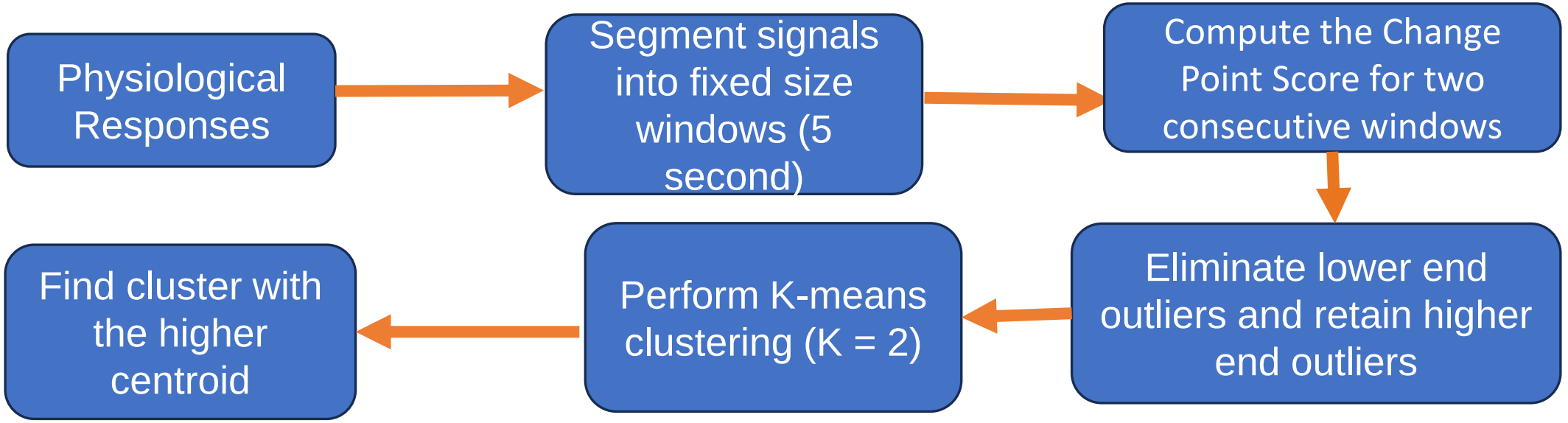
Labeling Opportune Moments



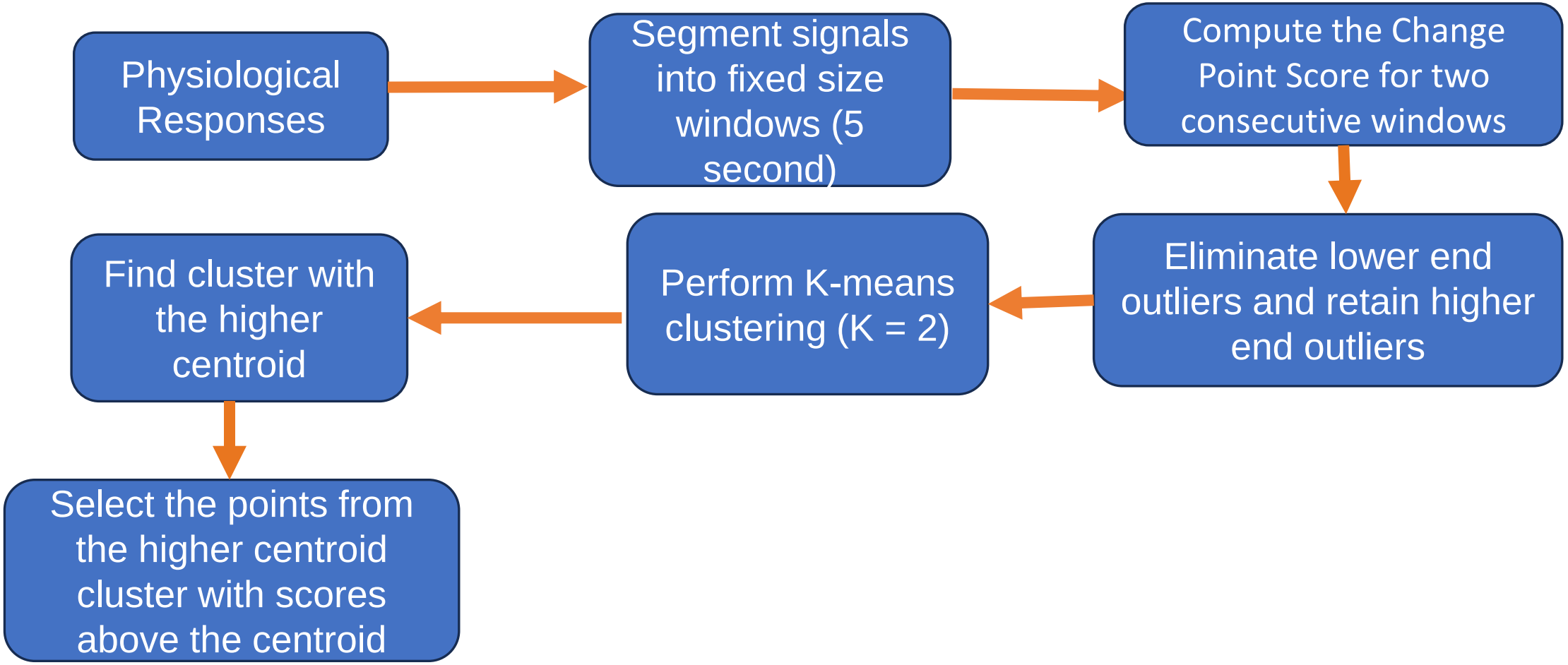
Labeling Opportune Moments



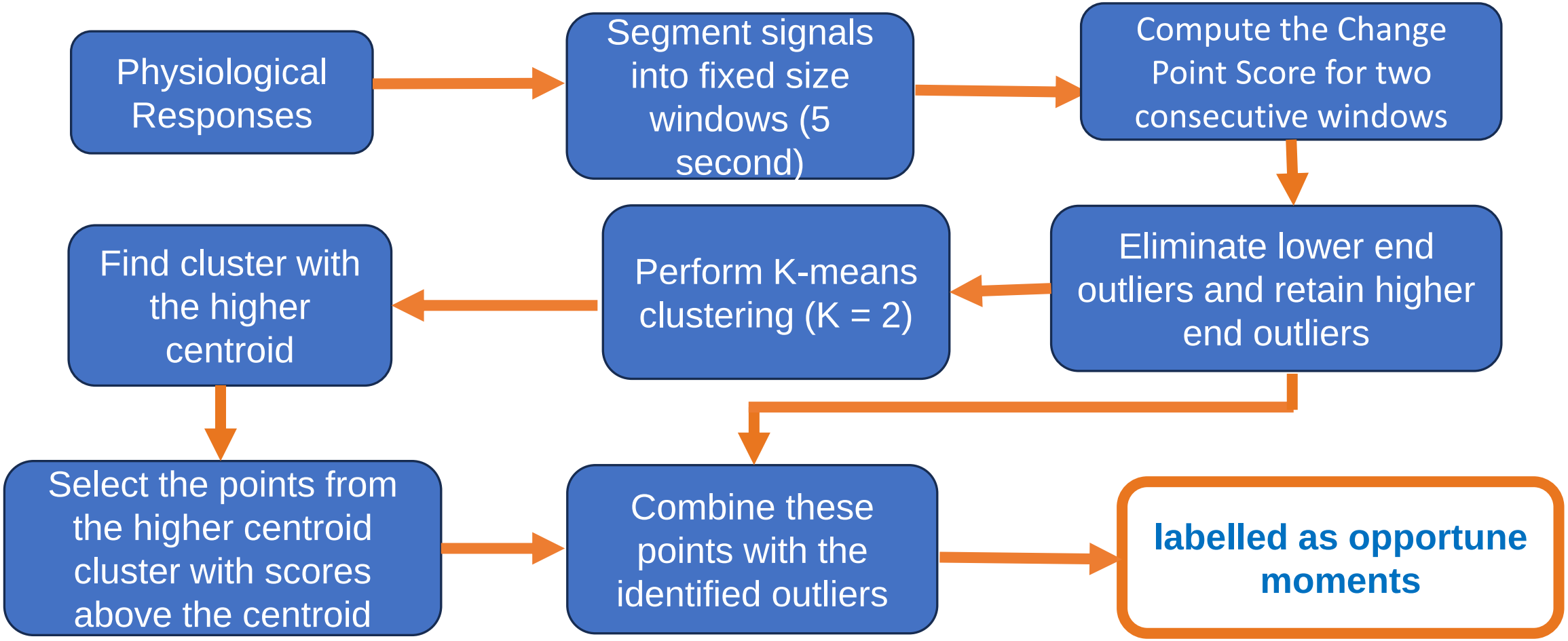
Labeling Opportune Moments



Labeling Opportune Moments



Labeling Opportune Moments



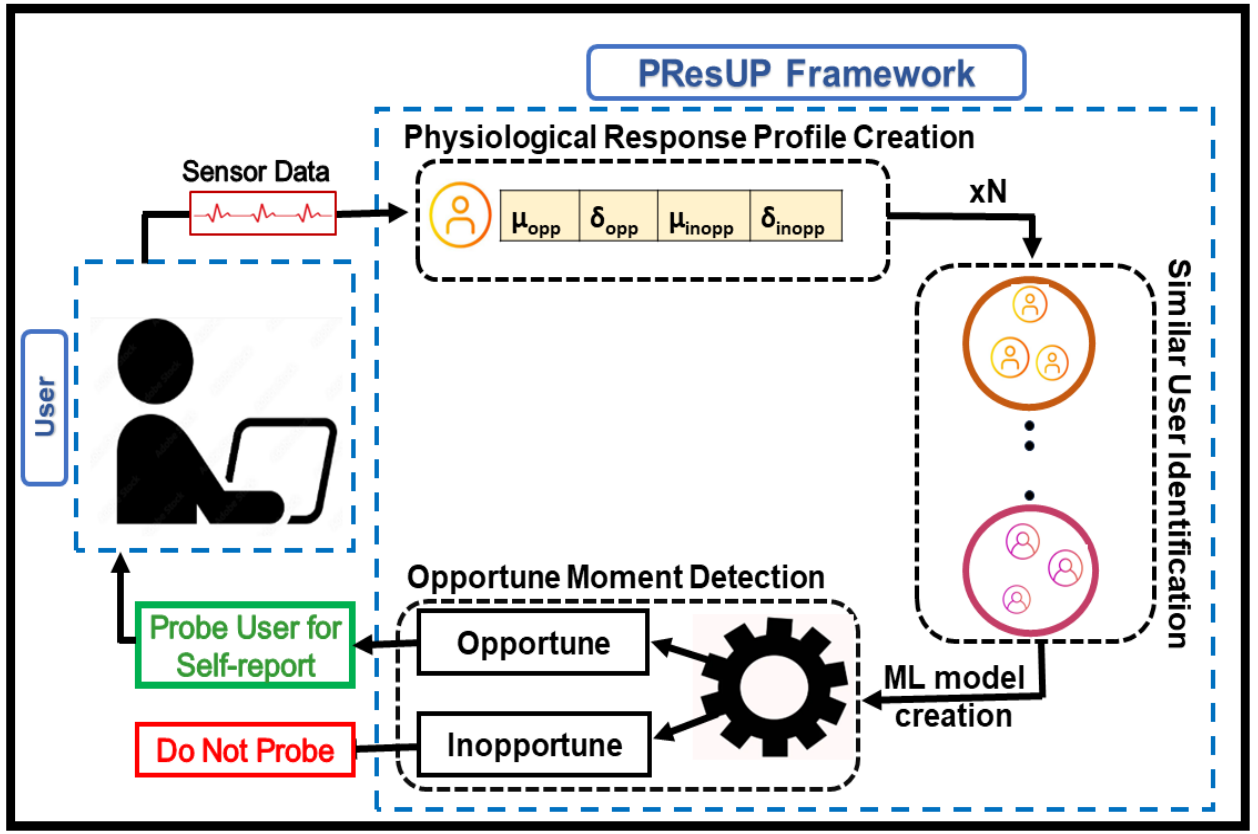
Data Description

- Total 7,290 segments.
- 11.6% are opportune and 88.4% are inopportune.

PResUP Framework

PResUP: Physiological Response based User Profiling

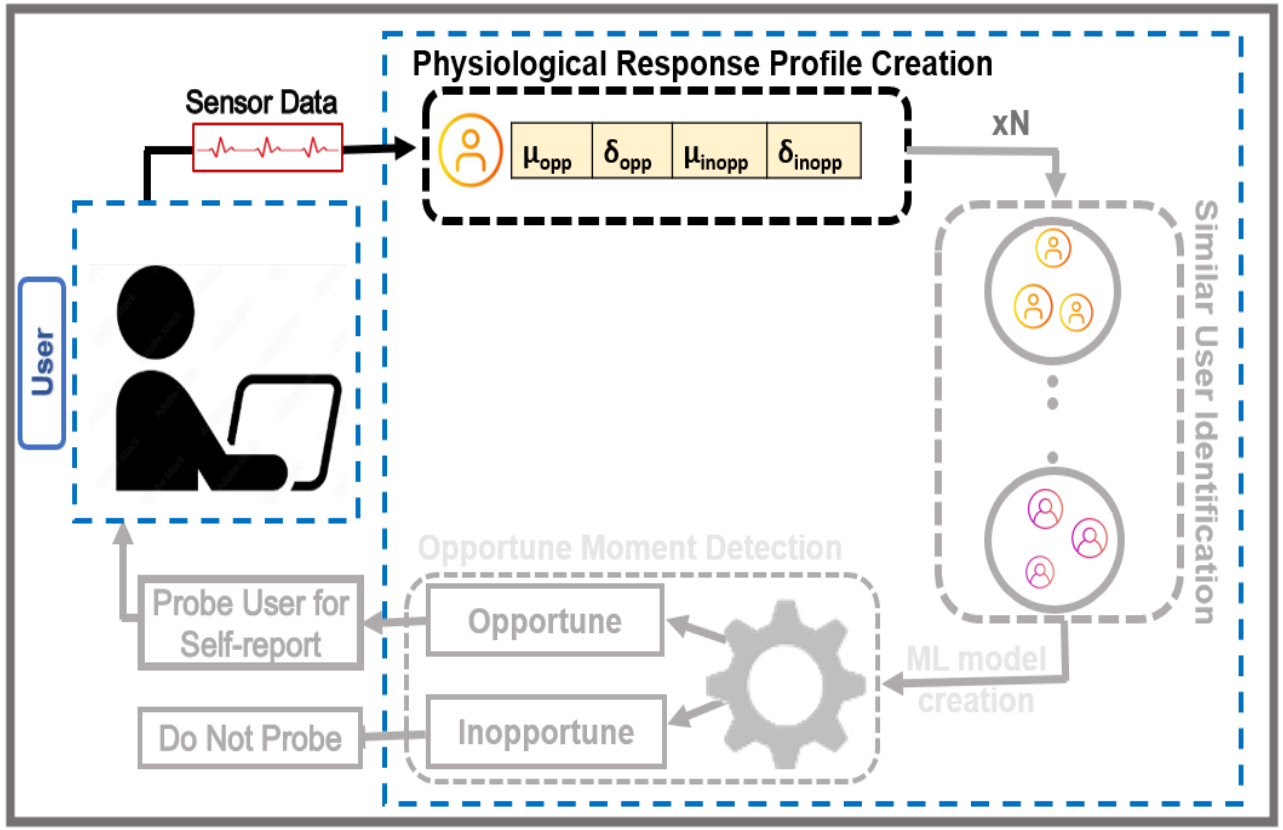
- Reduces continuous emotion self-report annotation effort
- Detects opportune moments



The **PResUP** framework

PResUP Framework

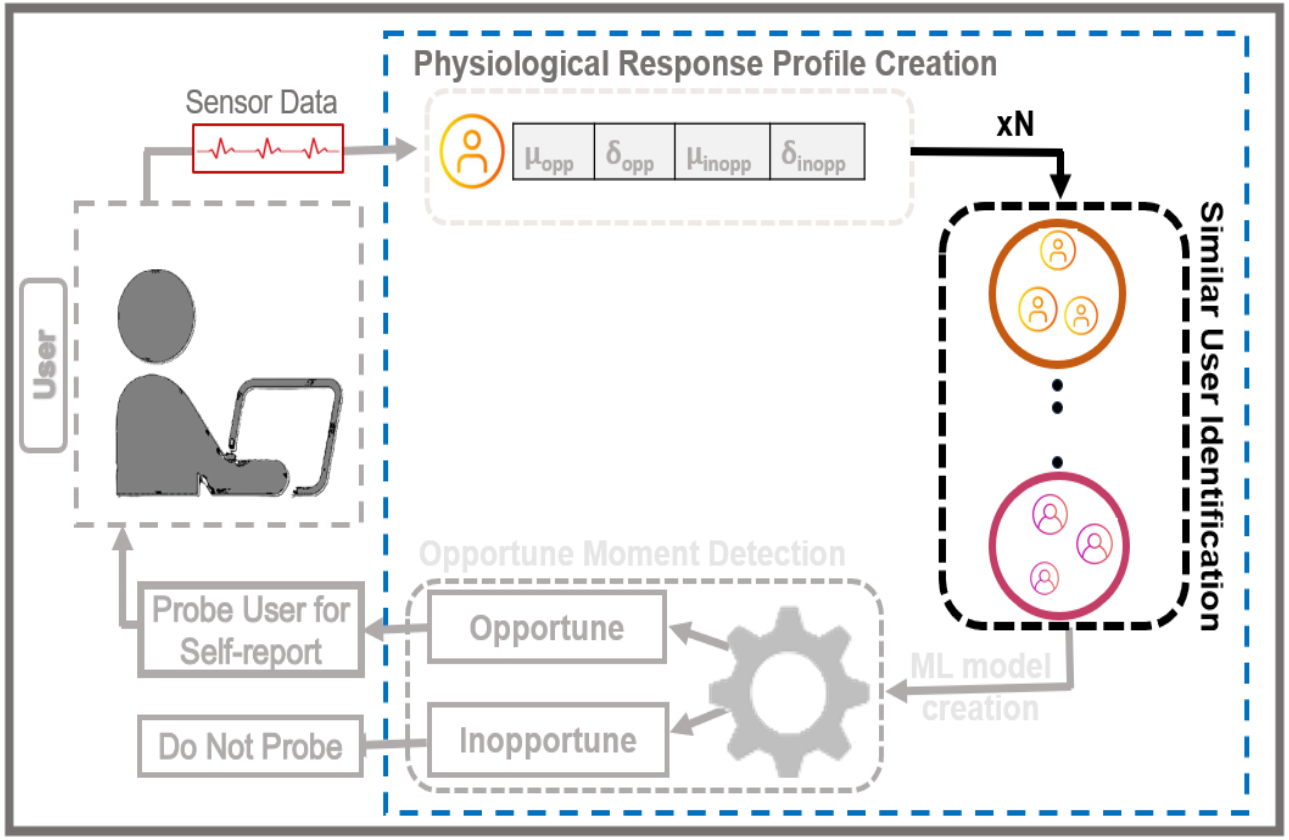
- Summarizes the statistical data of users' physiological responses during both opportune and inopportune moments.



The **PResUP** framework

PResUP Framework

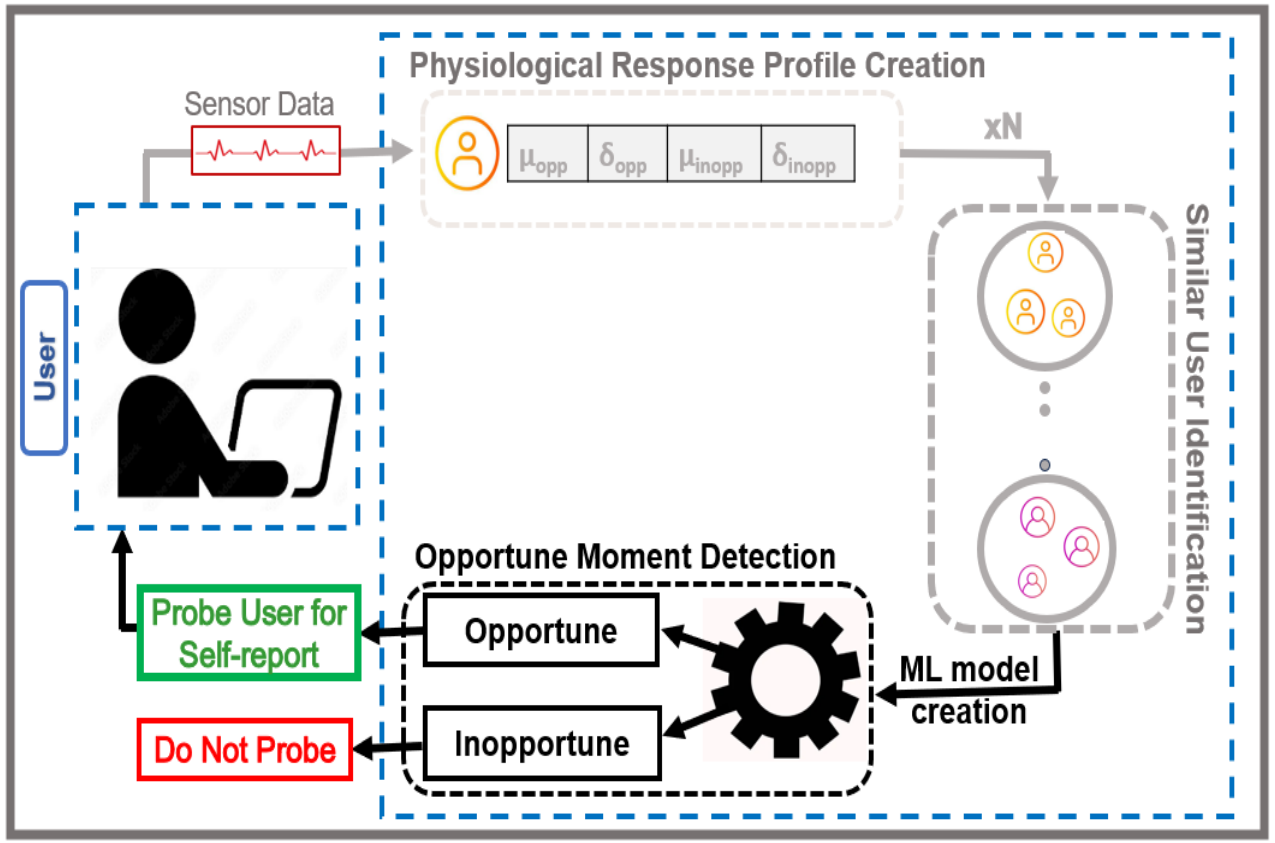
- Identify the group of users whose physiological response variations are similar.



The **PResUP** framework

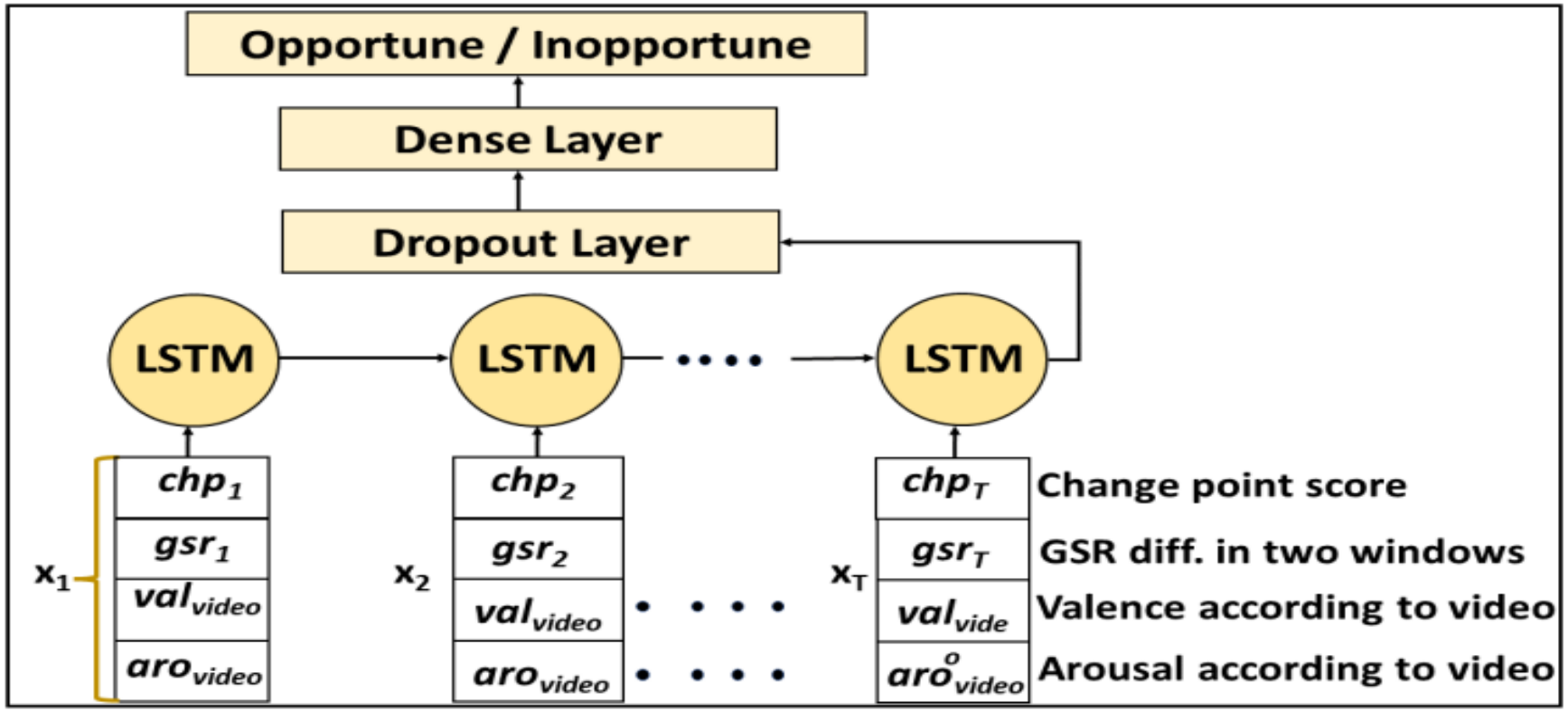
PResUP Framework

- Used LSTM model to detect the opportune moments



The **PResUP** framework

PResUP Framework



LSTM-based architecture used in the PResUP framework

Evaluation Metrices

- Probing Rate
- True Positives Rate (TPR)
- False Positives Rate (FPR)
- Likelihood Ratio (LR+) : $\frac{\text{TPR (Sensitivity)}}{\text{FPR}(1-\text{Specif icity})}$

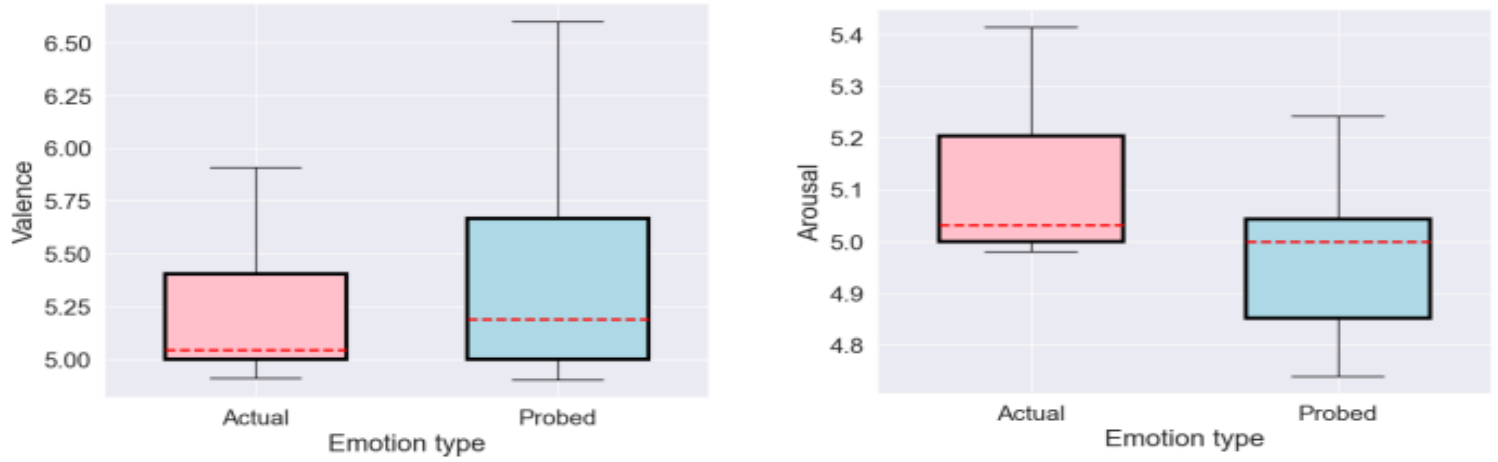
Evaluation

	Probing rate↓	TPR (%) ↑	FPR (%) ↓	LR+ ↑
TBS	30.38 (0.00)	100.00 (0.00)	100.00 (0.00)	1.00
PPS	20.38 (1.33)	69.21 (32.34)	16.16 (18.72)	4.28
FBS	6.50 (3.02)	70.49 (22.49)	15.02 (8.31)	4.69
GPS	6.37 (4.10)	81.30 (21.24)	14.93 (14.15)	5.45
APS	6.10 (4.04)	71.18 (20.86)	14.35 (15.98)	4.96
GBPS	6.07 (4.07)	82.12 (16.18)	13.68 (13.19)	6.00
RNNPS	6.01 (8.72)	82.66 (11.52)	12.88(5.24)	6.41
GRUPS	6.03 (8.98)	76.14 (10.75)	13.68 (5.37)	5.56
CNNPS	6.52 (6.76)	82.55(10.39)	15.47(6.96)	5.33
PResUP	**5.80 (2.89)**	**86.07 (11.64)**	**12.27 (9.67)**	**7.01**

Performance comparison of PResUP and baselines.

Evaluation

Actual : Original valence (and arousal) present in continuous annotation
Probed: Sampled valence (and arousal) using PResUP



(a) Comparison of valence scores **(b)** Comparison of arousal scores

No significant difference between ground truth continuous annotations and sampled values using PResUP framework

Dataset : K-EmoCon [1]

- 32 Participants (12F, 20M).
- 16 paired debates on accepting Yemeni refugees in Jeju, South Korea.
- Audio-visual recordings captured the debates.
- Continuous emotion annotations were made from three perspectives: subjects, partners, and external observers.

The participants sitting at a table for a debate

[1] C. Y. Park, N. Cha, S. Kang, A. Kim, A. H. Khandoker, L. Hadjileontiadis, A. Oh, Y. Jeong, and U. Lee, “K-emocon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversations,” Scientific Data, vol. 7, no. 1, pp. 1–16, 2020.

Dataset: Data Description

- Total 6644 segments.
- 13.6% are opportune and 86.3% are inopportune.

Evaluation

	Probing rate↓	TPR (%) ↑	FPR (%) ↓	LR+ ↑
TBS	237.28 (0.00)	100 (0.00)	100 (0.00)	1.00
PPS	74.46 (14.87)	81.26 (15.41)	13.01 (1.02)	6.25
FBS	32.32 (16.61)	80.22 (16.19)	4.04 (5.86)	19.86
GPS	33.89 (21.96)	75.65 (36.17)	5.16 (6.58)	14.66
APS	33.82 (16.41)	78.28 (17.51)	3.72 (5.74)	21.04
GBPS	33.78 (17.22)	80.01 (17.16)	3.74 (5.98)	21.39
RNNPS	34.35 (23.93)	68.39 (44.94)	6.56 (6.92)	10.42
GRUPS	40.96 (18.35)	50.51 (43.57)	11.60 (2.94)	4.35
CNNPS	44.5 (16.38)	73.53 (27.01)	11.41 (5.95)	6.44
PResUP	**31.36 (11.23)**	**82.26 (13.38)**	**3.41 (4.41)**	**24.12**

Performance comparison of PResUP and baselines.

Future scope

- Deployment of PResUP Framework
- Real-World Application
- Mobile Platform Applicability

Thank You!

Email: p20210016@goa.bits-pilani.ac.in

Website: https://sites.google.com/goa.bits-pilani.ac.in/swarnalibanik/

BITS Pilani
K K Birla Goa Campus