Assessment of Bipolar Disorder Using Heterogeneous Data of Smartphone-based Digital Phenotyping

[1]Hung-Yi Su, [1]Chung-Hsien Wu, [1]Cheng-Ray Liou, [2]Esther Ching-Lan Lin, [3]Po See Chen

[1]Department of Computer Science and Information Engineering, [2]Department of Nursing, [3]Department of Psychiatry
National Cheng Kung University, Tainan, Taiwan

National Cheng Kung University
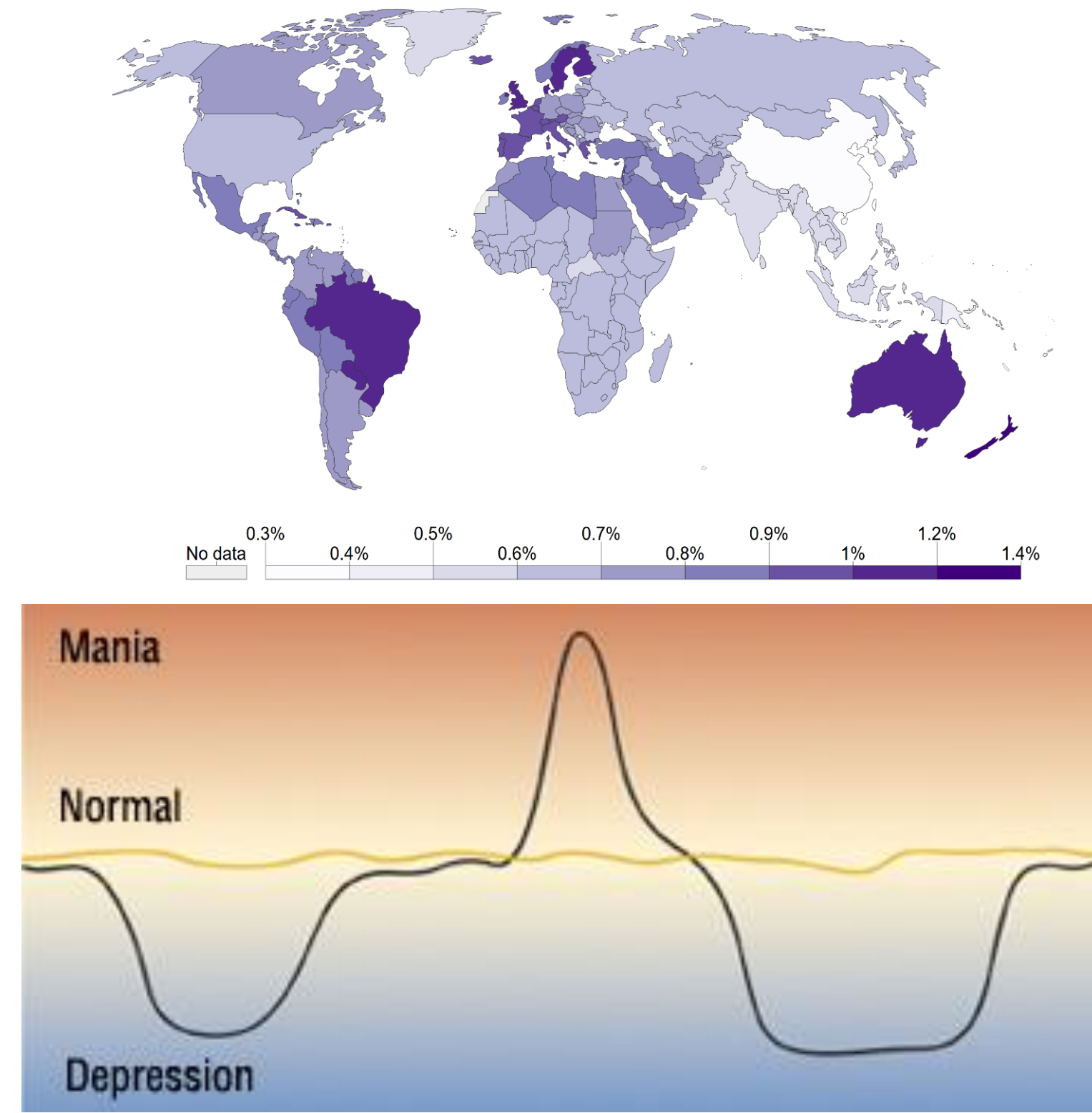
Poster Number : 1435

Introduction

➤ **Background:**
- Diagnosis of bipolar disorder(BD)
  - Judgment by professionally trained and experienced physicians.
  - Measurement using mental illness scales.
    - Hamilton Depression Rating Scale (HAM-D) and Young Mania Rating Scale (YMRS)
- Digital phenotypes

➤ **Problem and goal :**
- Missing data, Heterogeneous data.
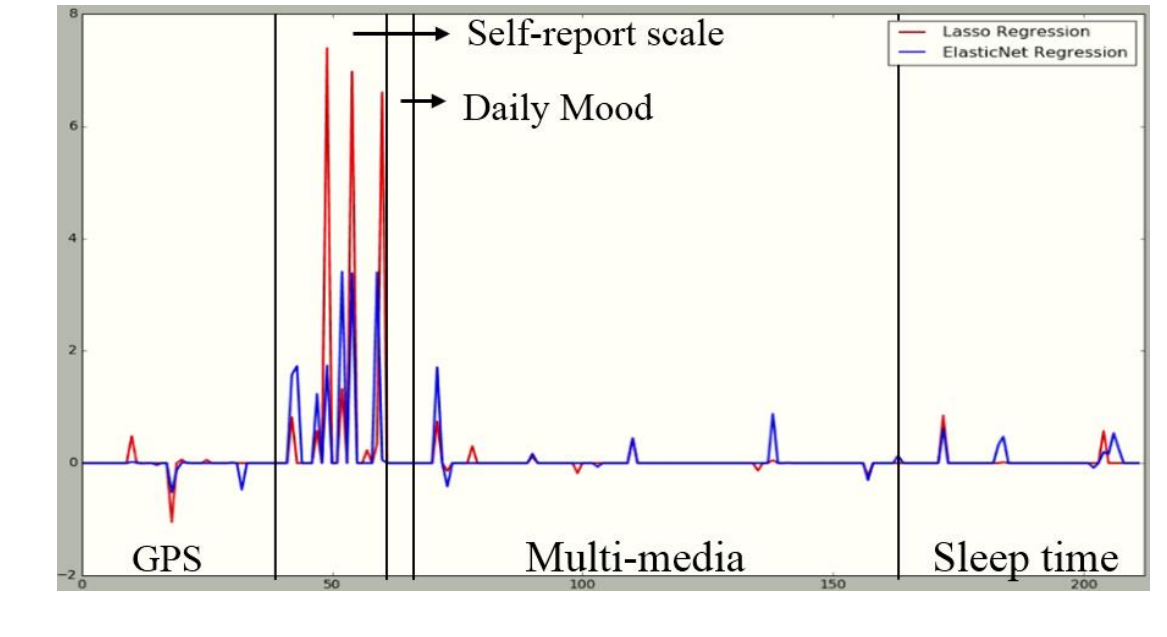- Convert digital footprints into behavioral characteristics and use them to predict BD.

Proposed Method

System Framework

Model pool

Feature extraction

Data	Functional
GPS	Entropy, Normalized Entropy, Home Stay, Location Variance, Transition Time, Total Distance, Number of Clusters, Diurnal Movement, Diurnal Movement on Normalized Coordinates, Diurnal Movement on the Distance from Home
Self-report scale	Score on each Question, Total Score
Daily mood	Mean, STD
Multi-media	Mean and STD of Seven class Subjective Emotion Profile, Mean and STD of Four class Objective Emotion Profile
Sleep time	Mean and STD of Sleep Duration Value, Mean and STD of Sleep Duration Class, Mean and STD of Sleep Midpoint Value, Mean and STD of Sleep Midpoint Class, Mean and STD of Sleep Regularity Value, Mean and STD of Sleep Regularity Class
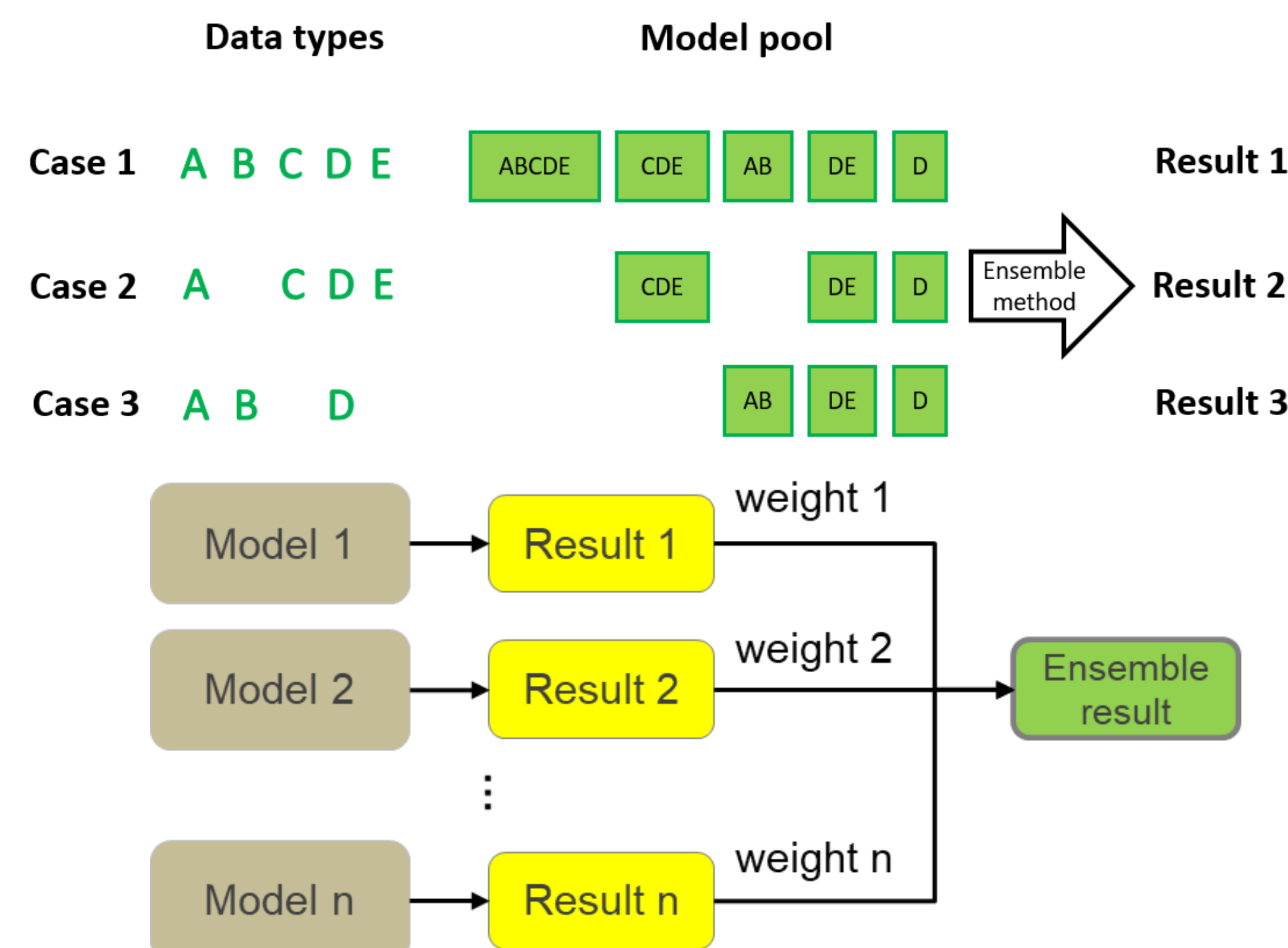
Feature Combination

Data combinations	Example
C_1^5	GPS, Daily mood, Daily mood, etc.
C_2^5	GPS + Daily mood, Daily mood + Self-report scale, etc.
C_3^5	GPS + Daily mood + Self-report scale, GPS + Daily mood + Sleep time, etc.
C_4^5	GPS + Daily mood + Self-report + Sleep time, GPS + Daily mood + Self-report + Multi-media, etc.
C_5^5	GPS + Daily mood + Self-report + Multi-media + Sleep time

Ensemble method

Experimental Results

Dataset

Bipolar Disorder Database

Subjects	84
Males and females	39, 45
Age (average)	38
HAM-D score	5.70
YMRS score	1.25

Model Performance Analysis

Model	HAM-D		YMRS	
	MAE	STD	MAE	STD
Lasso	2.87	1.2	1.06	0.5
ElasticNet	2.73	1.2	1.12	0.5
Polynomial (degree=2)	3.69	1.7	2.41	1.5
Polynomial (degree=3)	3.85	1.8	2.56	1.5
DNN	7.33	3.4	5.63	2.9

Heterogeneous Data Performance Analysis

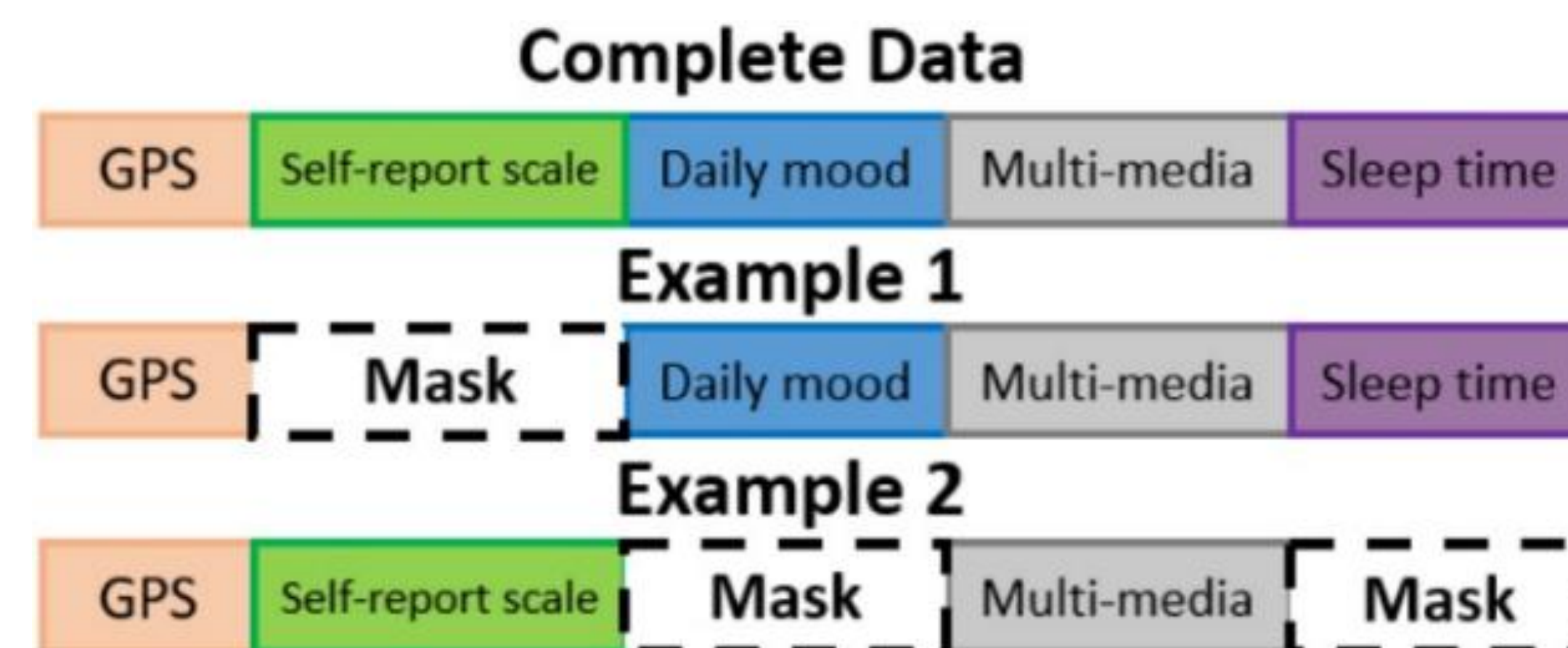| C_n^5	Lasso				ElasticNet			
	HAM-D		YMRS		HAM-D		YMRS	
	MAE	STD	MAE	STD	MAE	STD	MAE	STD
$n=1$	5.73	3.9	4.01	2.9	5.69	3.8	4.05	2.9
$n=2$	4.68	2.7	3.12	2.3	4.61	2.8	3.09	2.2
$n=3$	4.71	2.8	2.28	1.8	4.57	2.7	2.32	1.8
$n=4$	3.42	2.1	1.53	1.2	3.18	1.9	1.57	1.3
$n=5$	2.87	1.2	1.06	0.5	2.73	1.2	1.12	0.5

Feature Analysis of Heterogeneous Data

Number of Masks	Number of Features	HAM-D (MAE)	YMRS (MAE)
4	1	5.69	4.01
3	2	4.61	3.09
2	3	4.57	2.28
1	4	3.18	1.53
0	5	2.73	1.06

Performance of Ensemble Method

Conclusions

➤ Final experimental results (Mean value of ensemble method) :
- HAM-D : 4.15 (MAE)
- YMRS : 2.39 (MAE)

➤ Future work
- Collect more data samples
  - For more stable and effective training.
  - There may be much better choices and breakthrough in model selection.
- Add more data types

➤ With the data growing, the historical data of the subject can be used as a reference to establish a long-term tracking and personalized system.