Patient-Specific Modeling of Daily Activity Patterns for Unsupervised Detection of Psychotic and Non-Psychotic Relapses

2nd e-Prevention challenge: Psychotic and Non-Psychotic Relapse Detection using Wearable-Based Digital Phenotyping

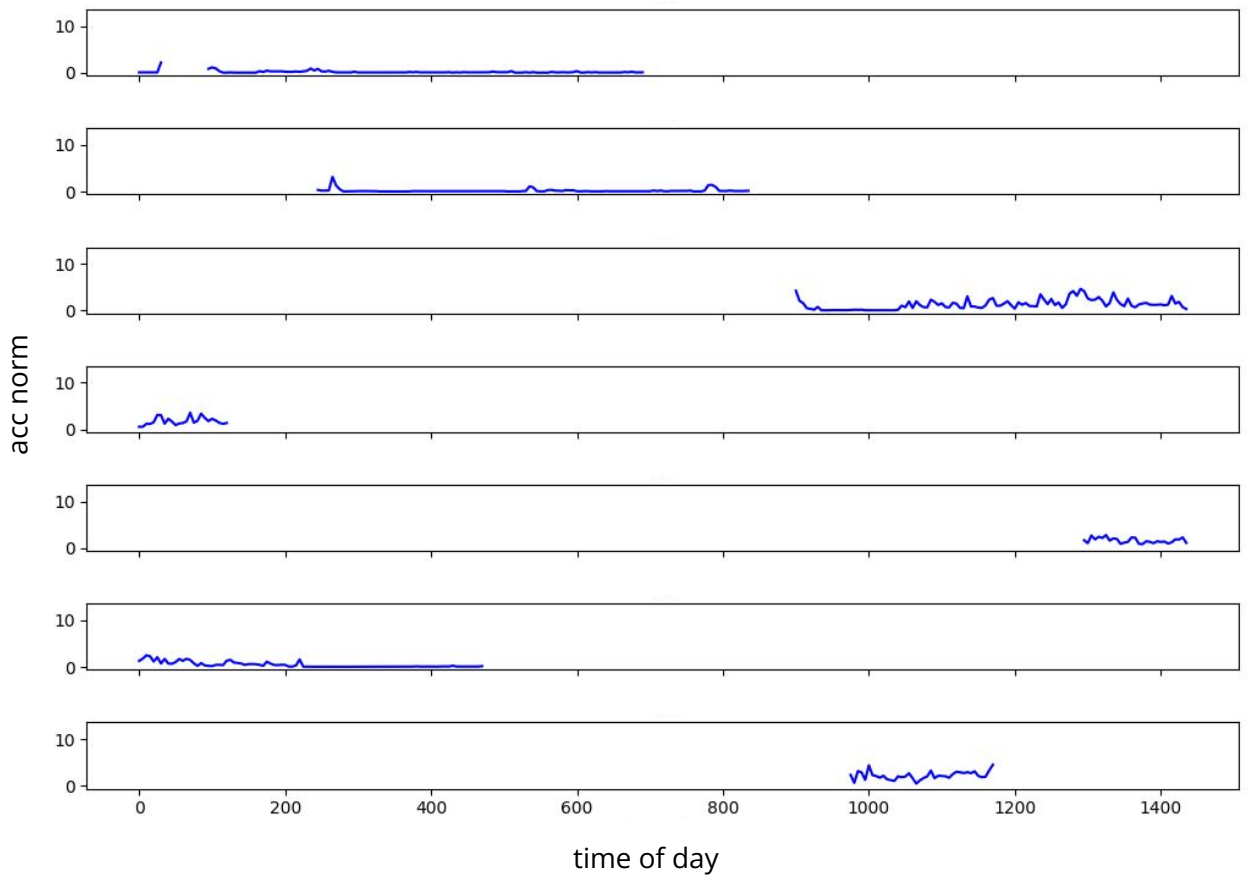
ICASSP 2024

Agenda

- Data Exploration
- Pre-Processing
- Model(s) and Training
- Outlier Detection
- Results

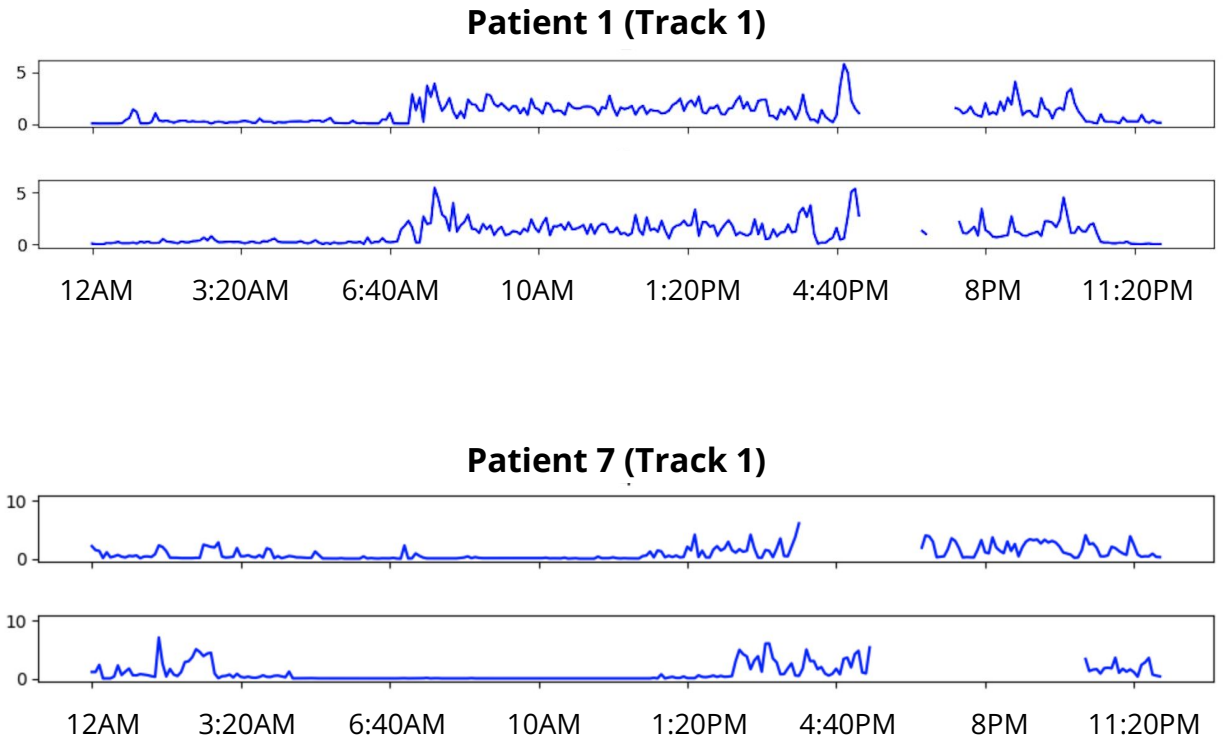
Data Exploration

- Missing data



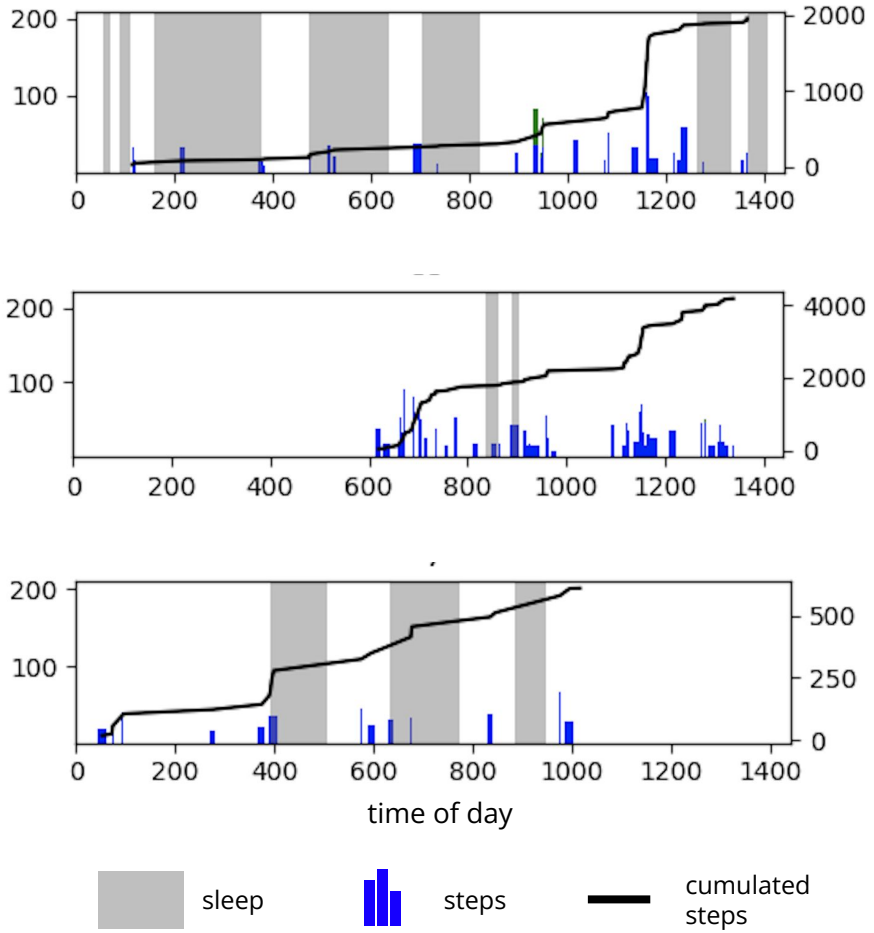
Data Exploration

- Missing data
- Differences in daily routines between patients



Data Exploration

- Missing data
- Differences in daily routines between patients
- Issues with sleep data



Pre-Processing

- Missing data
- Large differences between patients
- Issues with sleep data

Track 1: imputed missing values with median value for 5-minute time segment per patient

Track 2: discarded segments with missing values

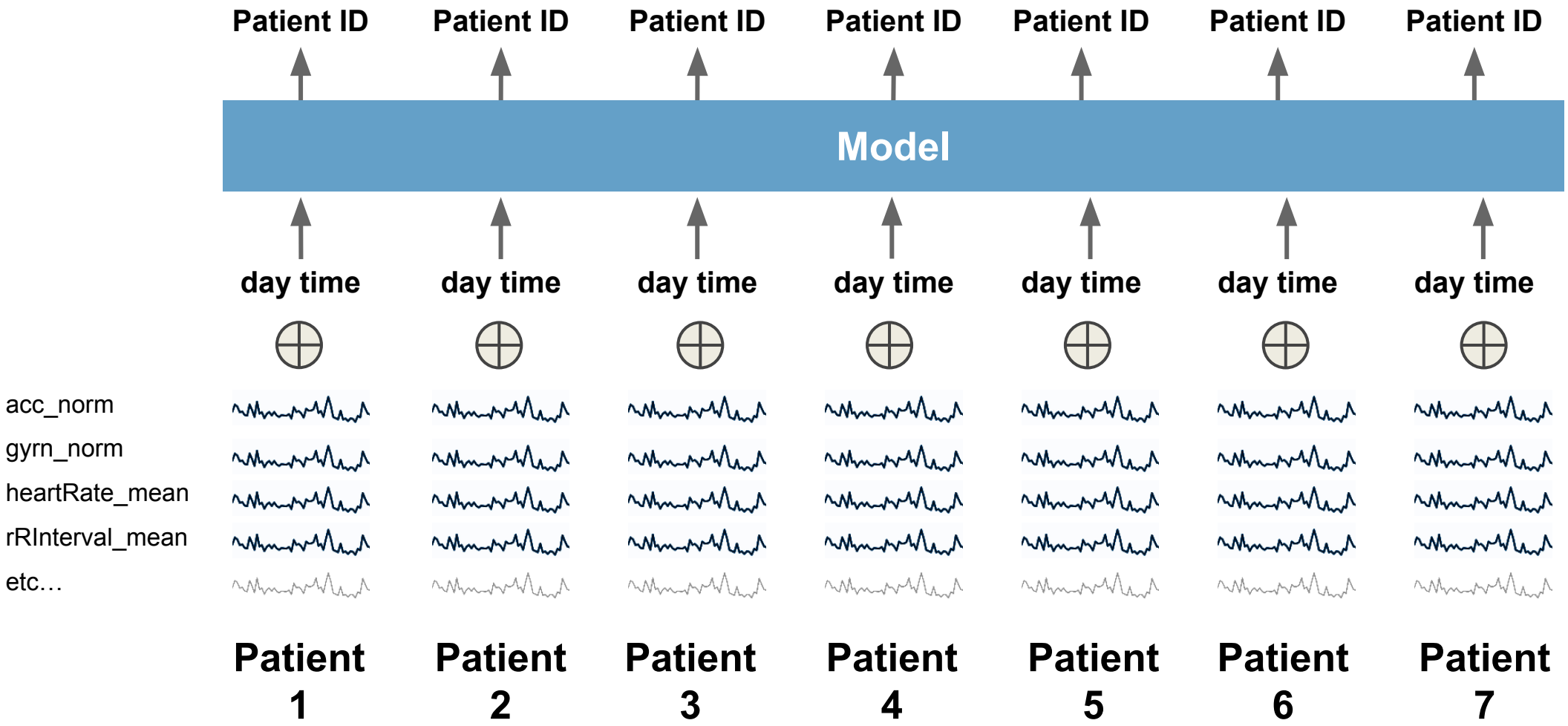
Pre-Processing

- Missing data
- Large differences between patients
- Issues with sleep data → dropped sleep data

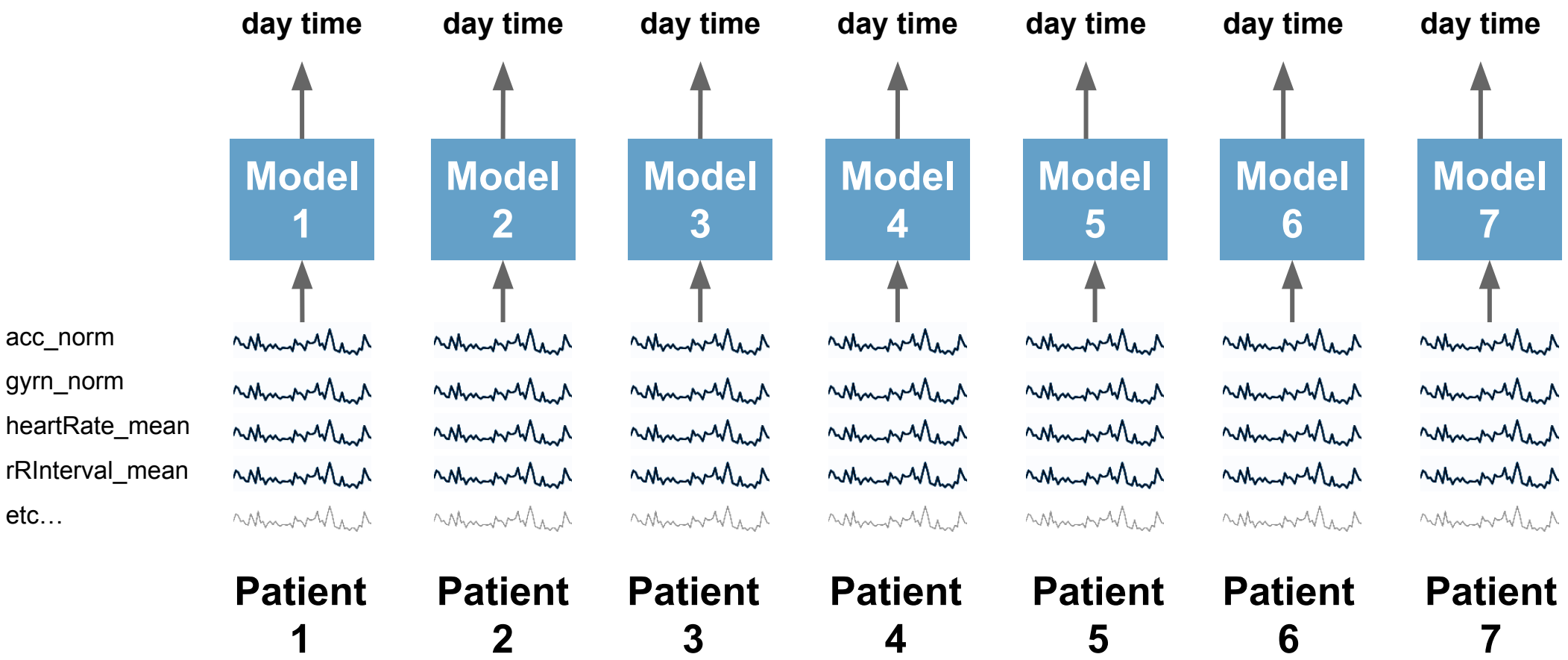
First Tries and Unsuccessful Ideas

- Echo-state networks
- Exponential Smoothing Transformers for Time-series Forecasting (ETSformer)

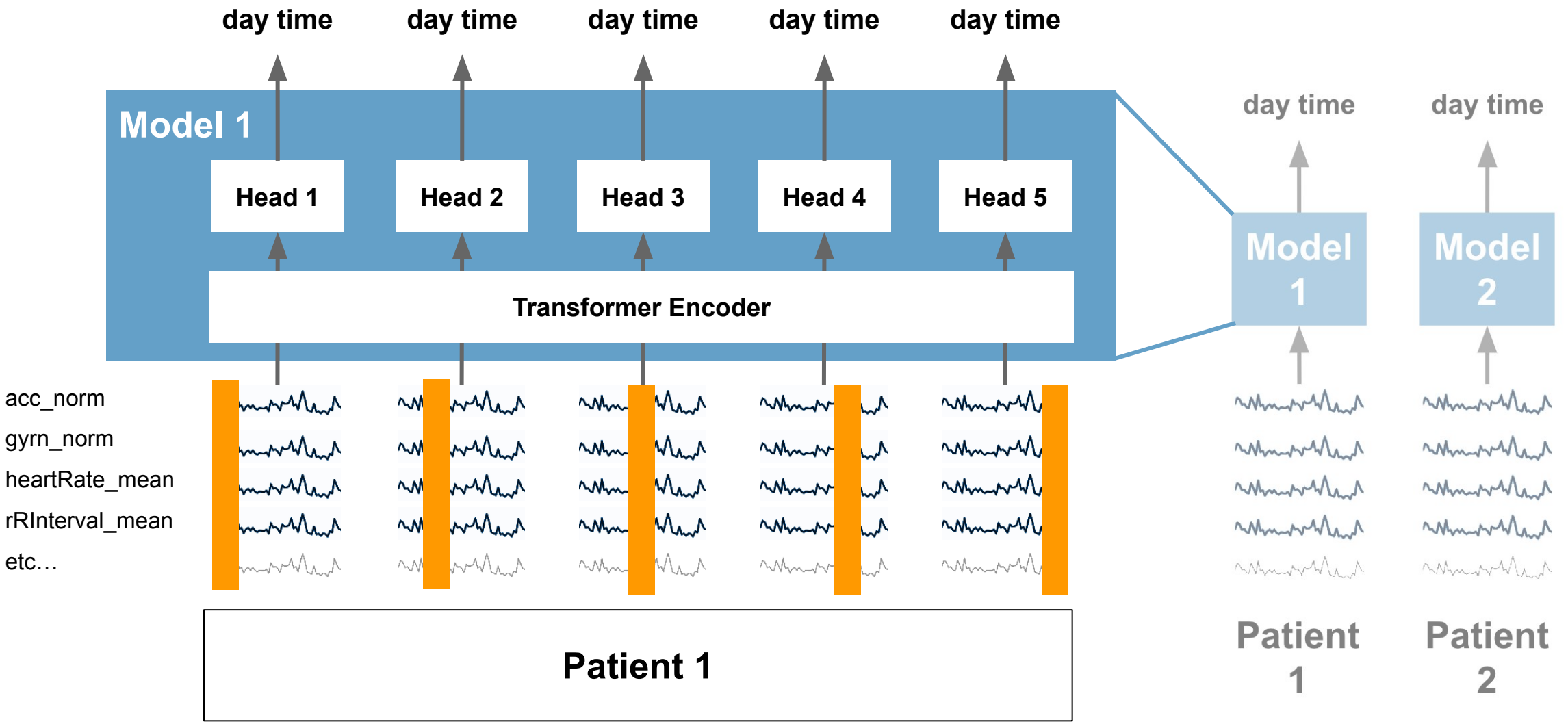
Baseline


Patient ID
Patient ID
Patient ID
Patient ID
Patient ID
Patient ID
Patient ID
Model
day time
day time
day time
day time
day time
day time
day time
acc_norm
gyrn_norm
heartRate_mean
rRInterval_mean
etc…
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Patient 7

Our Model: Track 1



Our Model: Track 2 (Ensemble)



Finding the Best Ensemble Approach

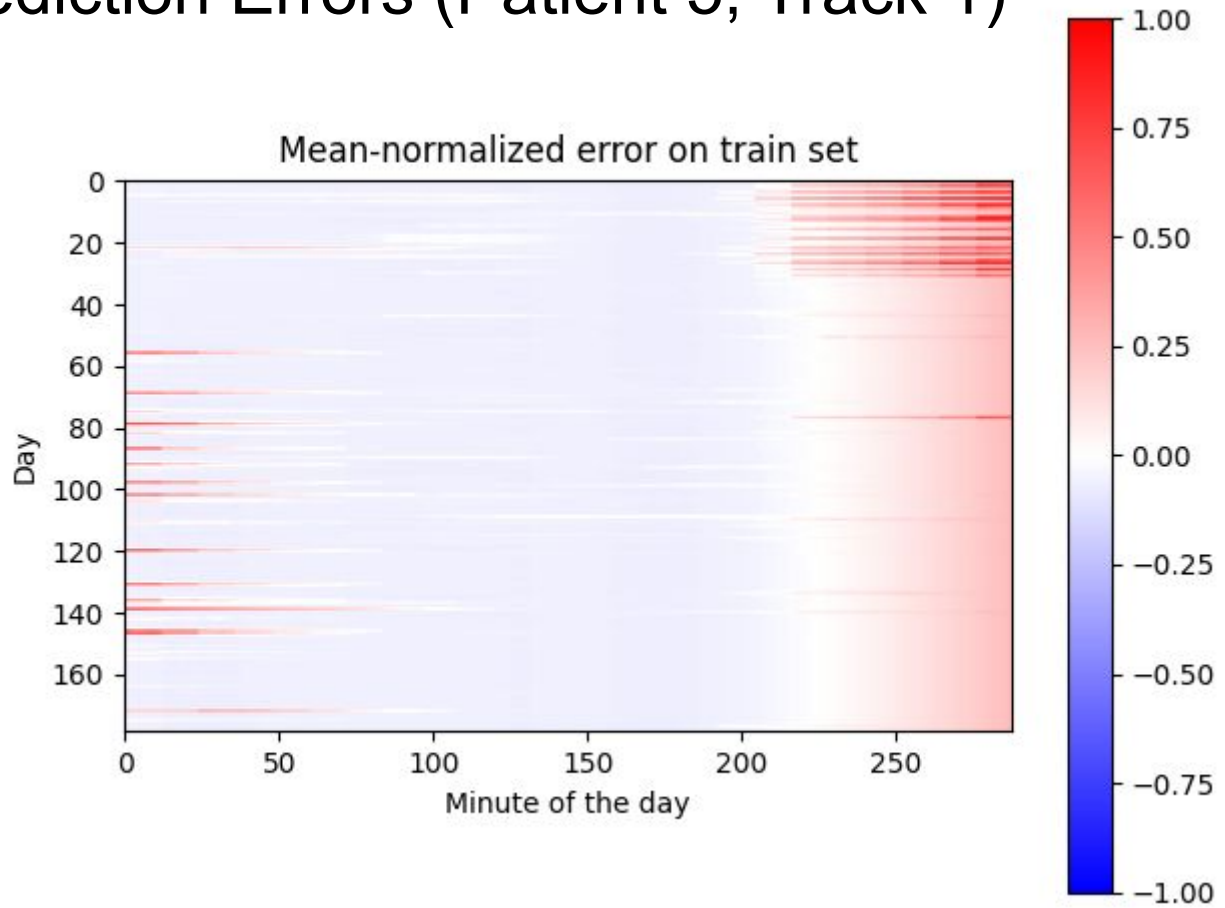
- Transformer encoder
  - shared for all patients but individual heads for each patient
  - patient-individual
- Ensemble heads
  - Linear
  - MLP
- Anomaly score determined by
  - variance of deterministic predictions
  - mean over covariance matrices of probabilistic predictions
  - comparison to mean over train set

Finding the Best Ensemble Approach

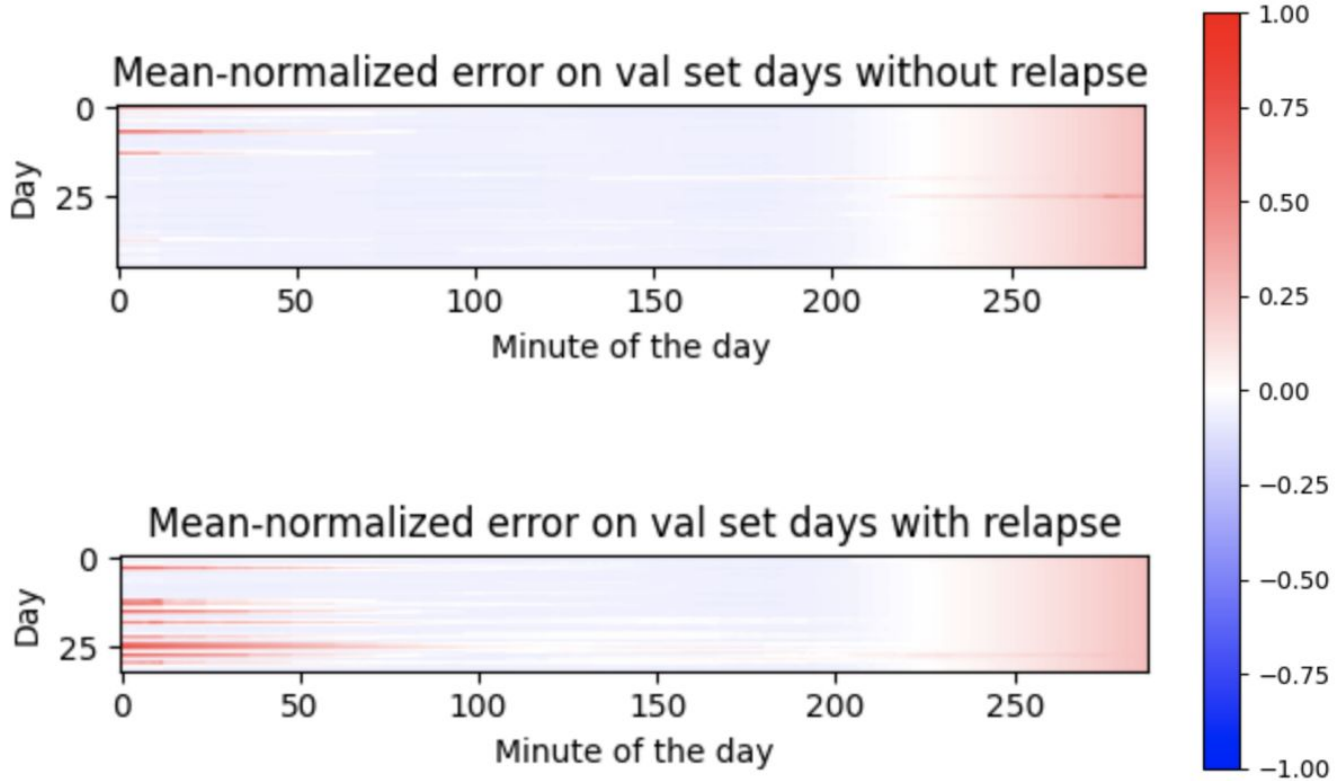
Final implementation

- Transformer encoder
  - shared for all patients but individual heads for each patient
  - patient-individual
- Ensemble heads
  - Linear
  - MLP
- Anomaly score determined by
  - variance of deterministic predictions
  - mean over covariance matrices of probabilistic predictions
  - comparison to mean over train set

Train set prediction Errors (Patient 5, Track 1)



Validation set prediction Errors (Patient 5, Track 1)



Other Changes

	Baseline			Ours	
	Track 1	**Track 2**		**Track 1**	**Track 2**
Sequence Length	24	48	→	72	24

Other Changes

	Baseline			Ours	
	Track 1	**Track 2**		**Track 1**	**Track 2**
Sequence Length	24	48	→	72	24
Output Dimension	32	32	→	64	64
Encoder Dimension	2048	2048		64	64

Other Changes

	Baseline			Ours	
	Track 1	**Track 2**		**Track 1**	**Track 2**
Sequence Length	24	48	→	72	24
Output Dimension	32	32		64	64
Encoder Dimension	2048	2048	→	64	64
Batch Size	16	16		64	16
Dropout	0.2	0.2	→	0.1	0.2
Epochs	10	10		100	50

Anomaly Detection: Track 1

1. Compute the mean prediction error for each day in training and validation sets

2. Calculate mean-normalized error for each day in validation set

$$e_{\text{norm}} = \frac{e_{\text{val}} - \bar{e}_{\text{train}}}{\max(e_{\text{train}}) - \min(e_{\text{train}})}$$

3. Classify as relapse day when mean-normalized prediction error is above zero

$$\text{score}(e_{\text{norm}}) = \begin{cases} 0 & \text{if } e_{\text{norm}} \leq 0 \\ 1 & \text{if } e_{\text{norm}} > 0. \end{cases}$$

Anomaly Detection: Track 2

1. Compute the mean prediction error for each day in training and validation sets

   a. Generate prediction for each member of ensemble p_i
   b. Calculate error based on squared Euclidean distance of each prediction to the prediction mean μ

$$e_{\text{val}} = \mathbb{E}_{\text{day}} \left[\mathbb{E}_i \left[\|p_i - \mu\|_2^2 \right] \right]$$

2. Calculate mean-normalized error for each day in validation set

$$e_{\text{norm}} = \frac{e_{\text{val}} - \bar{e}_{\text{train}}}{\max(e_{\text{train}}) - \min(e_{\text{train}})}$$

3. Classify as relapse day when mean-normalized prediction error is above zero

$$\text{score}(e_{\text{norm}}) = \begin{cases} 0 & \text{if } e_{\text{norm}} \leq 0 \\ 1 & \text{if } e_{\text{norm}} > 0. \end{cases}$$

Results

- **#3** for detecting non-psychotic relapses **(Track 1)**
- **#1** for detecting psychotic relapses **(Track 2)**

Track 1 Val	PR-AUC	ROC-AUC	AVG
Baseline	0.472	0.614	0.543
Ours	**0.680**	**0.665**	**0.672**
Track 2 Val	PR-AUC	ROC-AUC	AVG
Baseline	0.452	0.594	0.522
Ours	**0.694**	**0.669**	**0.681**
Track 1 Test	PR-AUC	ROC-AUC	AVG
Baseline	0.561	0.485	0.522
Ours	**0.595**	**0.574**	**0.584**
Track 2 Test	PR-AUC	ROC-AUC	AVG
Baseline	0.548	0.412	0.480
Ours	**0.563**	**0.444**	**0.504**

Thanks for your attention!

Code available at:
https://github.com/SvenGronauer/spgc-eprevention-icassp2024