# Personalised Anomaly Detectors and Prototypical Representations for Relapse Detection from Wearable-Based Digital Phenotyping


Adria Mallol-Ragolta[1,2], Anika Spiesberger[1,2],
Andreas Triantafyllopoulos[1,2], and Björn Schuller[1,2,3]

[1] CHI – Chair of Health Informatics, MRI, Technical University of Munich, Germany

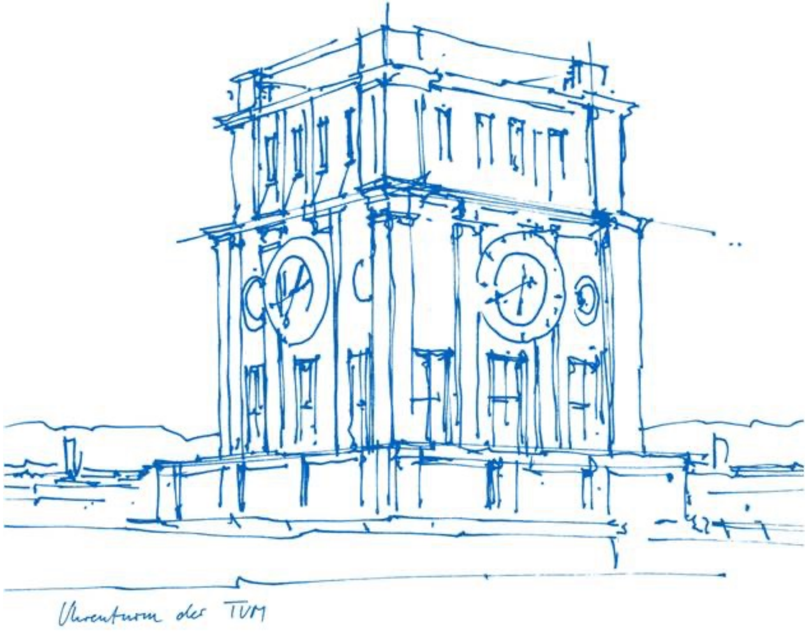
[2] EIHW – Chair of Embedded Intelligence for Health Care & Wellbeing, University of Augsburg, Germany

[3] GLAM – Group on Language, Audio, & Music, Imperial College London, UK



ICASSP 2024, 16.04.2024, Seoul (Korea)

Funded by Deutsche Forschungsgemeinschaft
German Research Foundation

# 1. Introduction

- The sensors embedded in current wearables allow capturing user-centred information
  - ➢ Characterisation of the user's digital phenotype
- The ubiquitous nature of wearables and their connectivity can generate large amounts of data
  - ➢ Artificial Intelligence (AI) techniques can favour its analysis
- Example problems exploiting wearable sensor data with AI:
  - – Person Identification [1]
  - – Human Activity Recognition [2]
- Research question

*Can we utilise wearable-based digital phenotyping for non-psychotic and psychotic relapse detection?*

[1] G. Retsinas, P. P. Filntisis, N. Efthymiou, E. Theodosis, A. Zlatintsi, and P. Maragos, "Person identification using deep convolutional neural networks on short-term signals from wearable sensors," in Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, Barcelona, Spain, 2020.

[2] A. Mallol-Ragolta, A. Semertzidou, M. Pateraki, and B. Schuller, "harAGE: A Novel Multimodal Smartwatch-based Dataset for Human Activity Recognition," in Proc. of the IEEE Intl. Conf. on Automatic Face and Gesture Recognition, Jodhpur, India, 2021.

# 1. Introduction

- Contribution to the 2nd e-Prevention Challenge [3]
  - Data collected with the e-Prevention system [4]
  - Patients' information gathered via a smartwatch and a tablet
- Investigated modalities
  - Gyroscope, accelerometer, and heart rate-related data
    - Captured continuously with the sensors embedded in a smartwatch
  - Sleep information
    - Reported on a daily basis
    - Shown indicative of relapses in patients with psychotic disorders [5]

---

[3] P. P. Filntisis, N. Efthymiou, G. Retsinas, A. Zlatintsi, C. Garoufis, T. Sounapoglou, P. Tsanakas, N. Smyrnis, and P. Maragos, "The 2nd e-prevention challenge: Psychotic and non-psychotic relapse detection using wearable-based digital phenotyping," in Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, Seoul, Korea, 2024.

[4] A. Zlatintsi, P. P. Filntisis, C. Garoufis, N. Efthymiou, P. Maragos, A. Menychtas, I. Maglogiannis, P. Tsanakas, T. Sounapoglou, E. Kalisperakis, T. Karantinos, M. Lazaridi, V. Garyfalli, A. Mantas, L. Mantonakis, and N. Smyrnis, "E-prevention: Advanced support system for monitoring and relapse prevention in patients with psychotic disorders analyzing long-term multimodal data from wearables and video captures," Sensors, vol. 22, 2022.

[5] F. Waite, N. Evans, E. Myers, H. Startup, R. Lister, A. G. Harvey, and D. Freeman, "The patient experience of sleep problems and their treatment in the context of current delusions and hallucinations," Psychology and Psychotherapy, vol. 89, 2015.

# 2. Methodology

- The e-Prevention Challenge consists of two tasks
  - Task 1: Wearable-based digital phenotyping for non-psychotic relapse detection
  - Task 2: Wearable-based digital phenotyping for psychotic relapse detection
- We investigate the two proposed tasks with the same approach
- We tackle the tasks as anomaly detection problems
  - Implementation of modality-dedicated autoencoders
  - Learning representative embeddings of the non-relapse training data

# 2. Methodology

## Smartwatch Sensor Data Pre-Processing

- Required step as noise and artifacts might be present in the data
- Sanity check of the raw measurements
  - ➢ Removing those outside modality-specific valid ranges
- Measurements segmentation
  - ➢ Containing 20 sec of continuous sensor measurements
- Measurements synchronisation
  - ➢ Gyroscope and accelerometer measurements sensed at 20 Hz
  - ➢ Heart rate-related information recorded at 5 Hz
  - ➢ Intermodality synchronisation looking for the closest segments in the time domain
    - – We discard those segments with an initial time difference greater than 0.2 sec
- We select 1/10 consecutive segments per day for training the autoencoders

# 2. Methodology

## Sleep Information Pre-Processing

- Night and day sleep differentiation
- For each sleep period, we extract:
  - Total sleep duration
  - Number of sleep intervals
  - Duration of the longest interval
- In total, we characterise the daily sleep information with six features

# 2. Methodology

## Relapse Detection Models

- We utilise autoencoders to learn embedded representations of the sensor measurements
- Each modality is modelled with its own autoencoder
    - ➢ The autoencoders corresponding to the smartwatch sensor data are based on GRU-RNN
    - ➢ The sleep autoencoder implements fully connected layers

- Personalised anomaly detection
    - ➢ Operates on the embeddings space of each patient individually
    - ➢ Computes a prototypical representation of the training, non-relapse embeddings
        - – Characterised by their mean and standard deviation
    - ➢ Models the personal distribution of embeddings with an Elliptic Envelope

# 2. Methodology

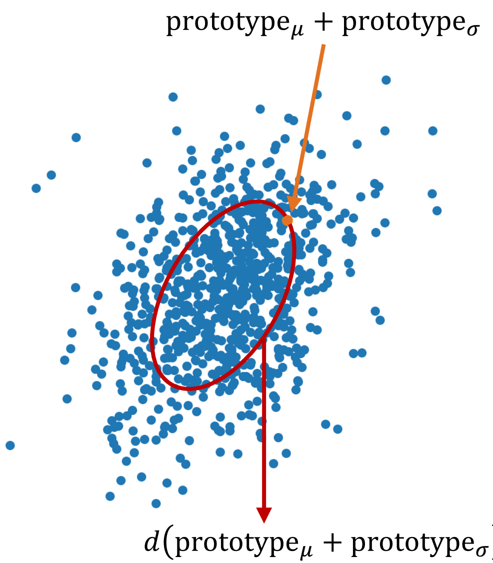
## Relapse Detection Models

- At validation and test times
  - ➢ The likelihood $p$ that the current segment corresponds to a relapse state is defined as:

$$p = \frac{d(\text{embedding})}{2 * d(\text{prototype}_{\mu} + \text{prototype}_{\sigma})}$$

  - ➢ The function $d(\cdot)$ corresponds to the Mahalanobis distance of the input representation in the training, non-relapse distribution

- Overall daily score computed as the average among the daily segment predictions




** Point cloud distribution created for demonstration purposes only*

# 3. Experimental Results

- Performance comparison of unimodal and multimodal systems
  - ➢ Resulting from all the possible combinations among the four considered modalities
- Models evaluation metric
  - ➢ Macro-average (AVG) of the AUROC and AUPRC scores over the daily predictions

$$\mathrm{AVG} = \frac{\mathrm{AUROC} + \mathrm{AUPRC}}{2}$$

*AUROC*: Area Under the Receiver Operating Characteristic Curve

*AUPRC*: Area Under the Precision-Recall Curve

# 3. Experimental Results

## Summary of the results (selected) in the non-psychotic relapse detection track (Track 1)

**Table 1**. Unimodal and multimodal systems combining S(leep) and A(ccelerometer) information. Results reported in percentage (%).

| Modality | Validation Set | | | Test Set | | |
|---|---|---|---|---|---|---|
| | AUROC | AUPRC | AVG | AUROC | AUPRC | AVG |
| S | 48.2 | 59.3 | 53.7 | 58.0 | 55.5 | 56.7 |
| S⊕A | 48.1 | 57.3 | 52.7 | 55.2 | 54.3 | 54.7 |
| Baseline | 61.4 | 47.2 | 54.3 | 56.1 | 48.5 | 52.3 |
| Random Chance | 50.0 | 32.6 | 41.3 | 50.0 | 43.0 | 46.5 |

# 3. Experimental Results

## Summary of the results (selected) in the psychotic relapse detection track (Track 2)

**Table 2**. Unimodal and multimodal systems combining S(leep), H(eart rate)-related and A(ccelerometer) information. Results reported in percentage (%).

| Modality | Validation Set | | | Test Set | | |
|---|---|---|---|---|---|---|
| | AUROC | AUPRC | AVG | AUROC | AUPRC | AVG |
| $S \oplus A$ | 52.9 | 49.7 | 51.3 | 49.5 | 48.6 | 49.1 |
| $S \oplus H \oplus A$ | 51.6 | 50.4 | 51.0 | 49.3 | 50.5 | **49.9** |
| Baseline | 59.4 | 45.2 | 52.2 | 54.8 | 41.2 | 48.0 |
| Random Chance | 50.0 | 34.9 | 42.4 | 50.0 | 34.7 | 42.4 |

# 4. Conclusions – Take Home Messages

- We used autoencoders to learn representative embeddings of the non-relapse training data, corresponding to gyroscope, accelerometer, heart rate-related measurements, and sleep information
- We determined the likelihood of relapse at inference time by computing the Mahalanobis distance between the embedded representations extracted from the unseen data and the training distribution
- For Track 1, the unimodal system exploiting the sleep information obtained the highest score, 56.7 %
- For Track 2, the multimodal system combining the embedded representations from the sleep, heart rate-related, and accelerometer modalities scored the best performance, 49.9 %
    - ➢ This system obtained the 2nd best performance of the challenge

## Future Works

- Personalisation of the anomaly detection thresholds
- Investigation of architectures and methods to extract effective embeddings from time series data