HUMAN DATA ANALYTICS

Deep Learning Techniques for Gesture Recognition: Dealing With Inactivity

Matteo Drago, Riccardo Lincetto
Dept. of Information Engineering, University of Padova, Italy

Prof. Michele Rossi

matteo.drago@studenti.unipd.it
riccardo.lincetto@studenti.unipd.it

September, 5, 2018

Outline

- Introduction
- Dataset description
- Related Work
- Our proposal
- Discussion of results
- Conclusion and future works

Introduction

Human Activity Recognition:

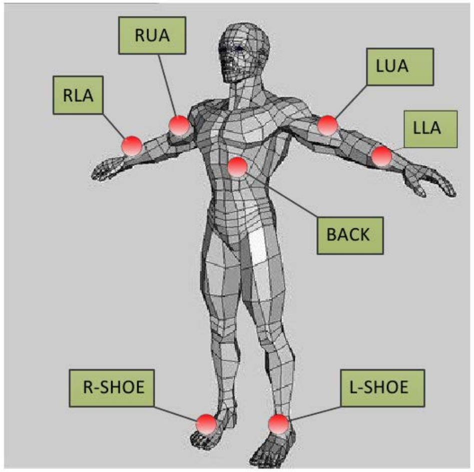
- Visual detection from images and video
- Gesture recognition from sensor-based data

In the past decade, many models have been designed for time series classification.

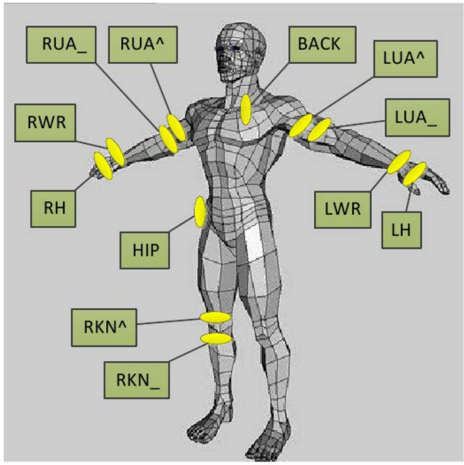
Main problems:

- LACK of BENCHMARKING DATASET to compare different solutions
- ABSENCE of DETAILS in most of the models presented in the literature

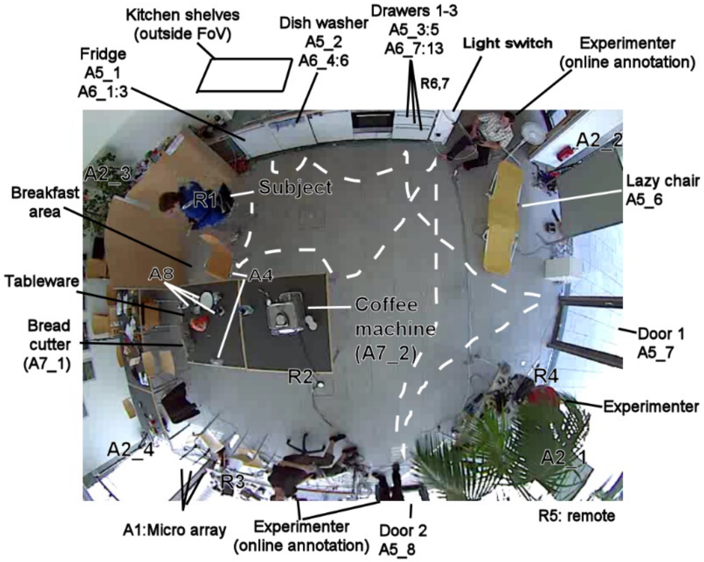
OPPORTUNITY DATASET



= Complete Inertial Measurement Unit



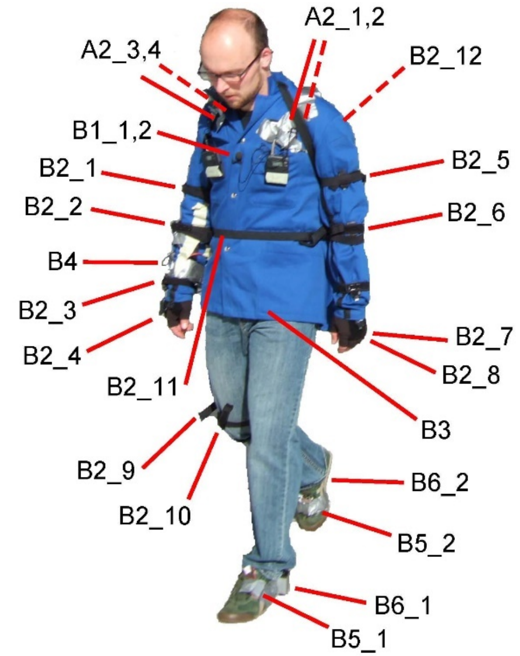
= Triaxial Accelerometer



[1] R. Chavarriaga, H. Sagha, A. Calatroni, S. T. Digumarti, G. Tröster, J. del R. Millán, and D. Roggen, "The opportunity challenge: A benchmark database for on-body sensor-based activity recognition," *Pattern Recognition Letters*, 2013.

HUMAN DATA ANALYTICS

- 4 different subjects
- 7 Inertial Measurement Units
- 12 accelerometer sensors



113 channels of measurements

Data has been collected in two distinct modalities :

- 5 sessions of **Activity of Daily Living** (ADL)
- **Drill** : 20 repetitions of low level activities

Multiclass Classification Problem

TASK A:

- Classification of high level gestures / modes of locomotion

Standing, Walking, Lying, Sitting

TASK B2:

- Recognition of low level gestures (17 in total)

Open Dishwasher, Close Dishwasher, Open Fridge, Close Fridge,

Open Drawer 1, Close Drawer 1, Open Door 1, Close Door 1, …

Both tasks comprehend the *Null Class*, which represents inactivity.
A more detailed discussion on this in a couple of slides

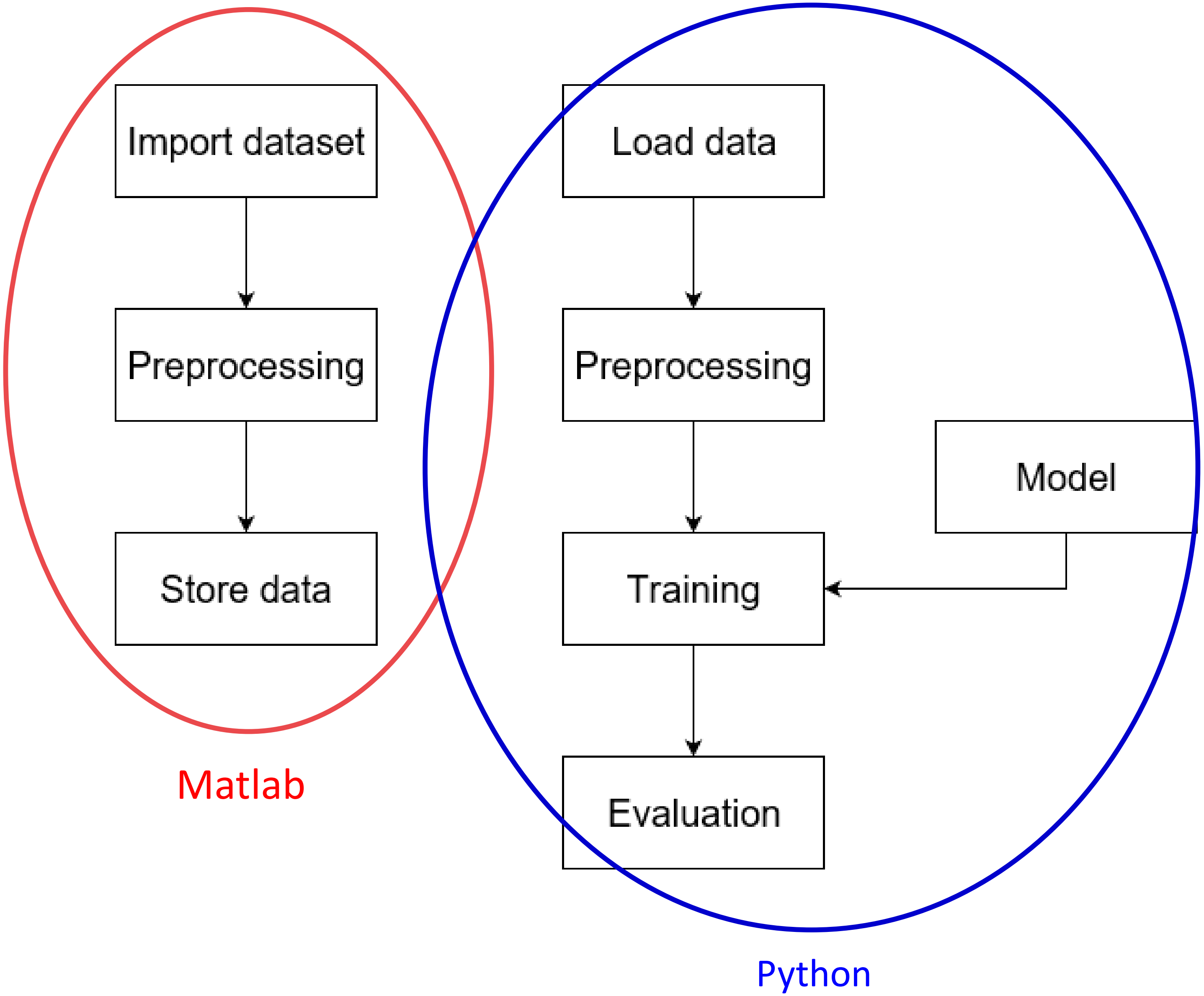
State of the Art

In the literature there’s no shortage of models trying to solve the problem. For example:

- Complete framework in [2] that performs interpolation in preprocessing and tries to cope with class imbalances (**1NN** and **SVM** as classifiers)
- **Convolutional layers** were implemented in NN along with **ReLU** and **pooling** layers in [3], in order to extract better features; they also segment the datasets into windows of samples
- A complete comparison can be found in [4] were they also implement a model comprehensive of both **convolutional** and **LSTM** layers, in order to extract sensible features and exploit the correlation among consecutive samples and independent windows

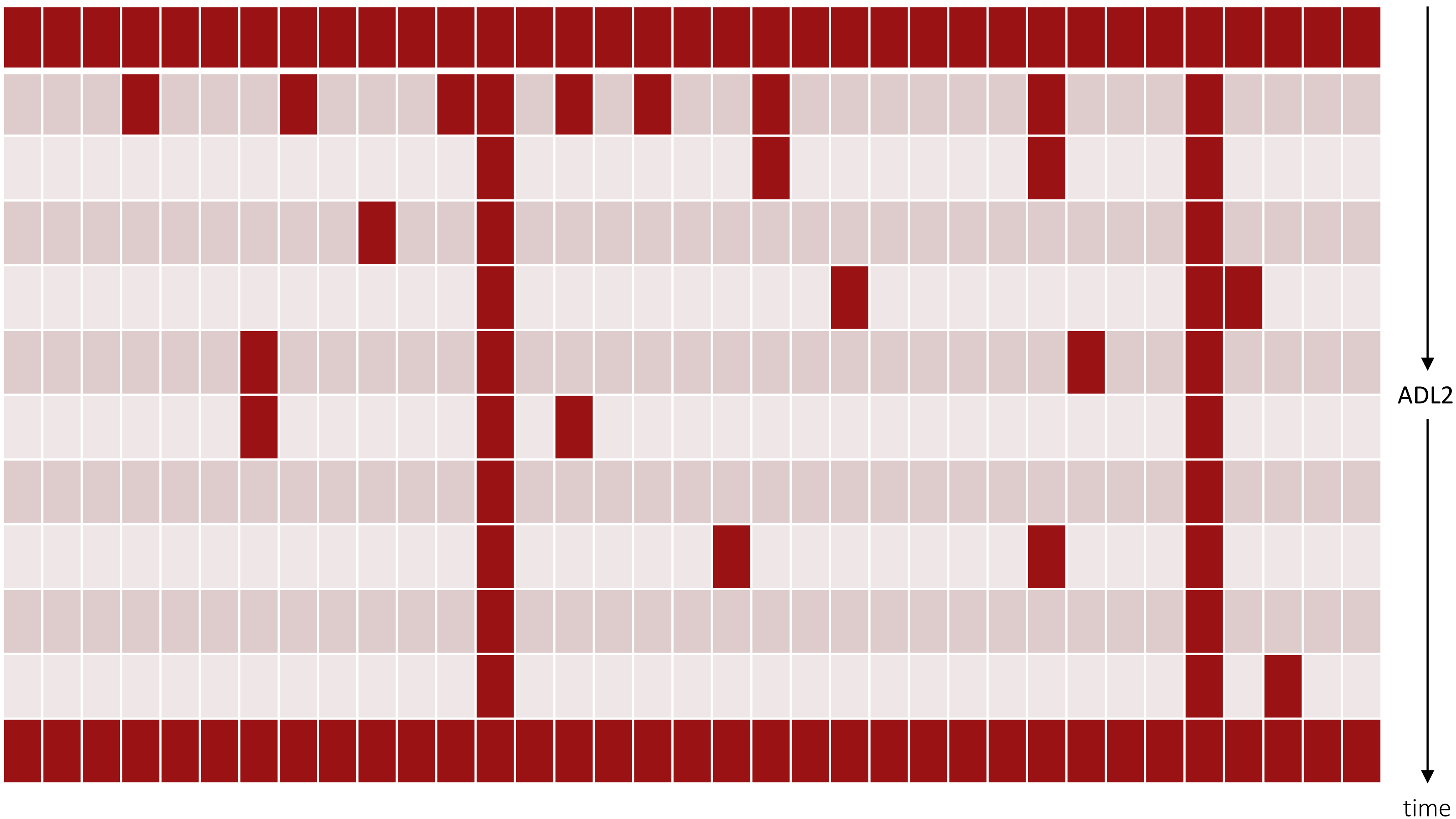
[2] H. Cao, M. N. Nguyen, C. Phua, S. Krishnaswamy, and X. Li, “An integrated framework for human activity classification.,” in *UbiComp*,pp. 331–340, 2012.
[3] J. Yang, M. N. Nguyen, P. P. San, X. Li, and S. Krishnaswamy, “Deep convolutional neural networks on multichannel time series for human activity recognition.,” in *Ijcai*, vol. 15, pp. 3995–4001, 2015.
[4] F. Li, K. Shirahama, M. A. Nisar, L. Köping, and M. Grzegorzek, “Comparison of feature learning methods for human activity recognition using wearable sensors,” *Sensors*, vol. 18, no. 2, p. 679, 2018.

Processing Pipeline of Our Proposal



Preprocessing

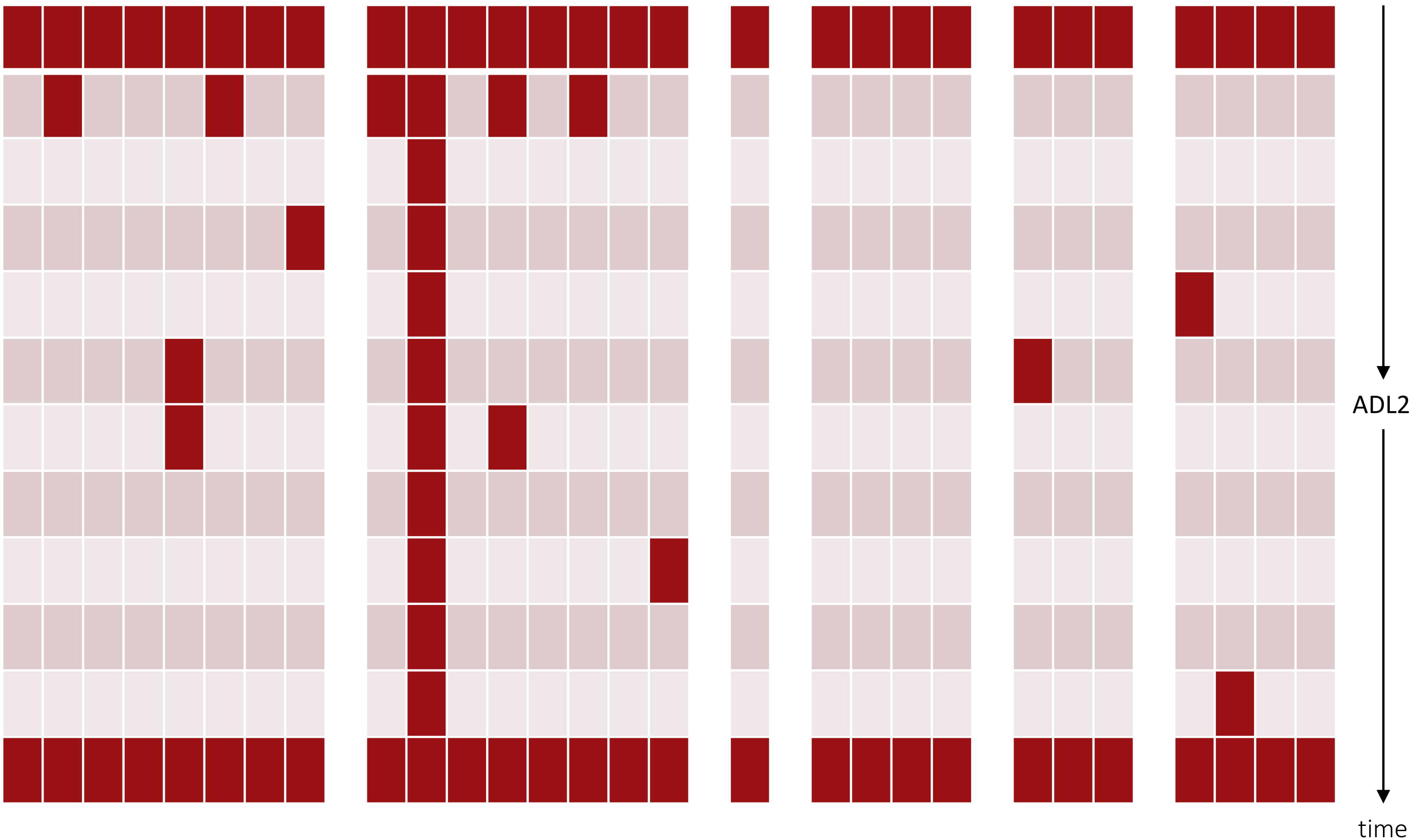
Original dataset (Dark squares = NaN values)



HUMAN DATA ANALYTICS

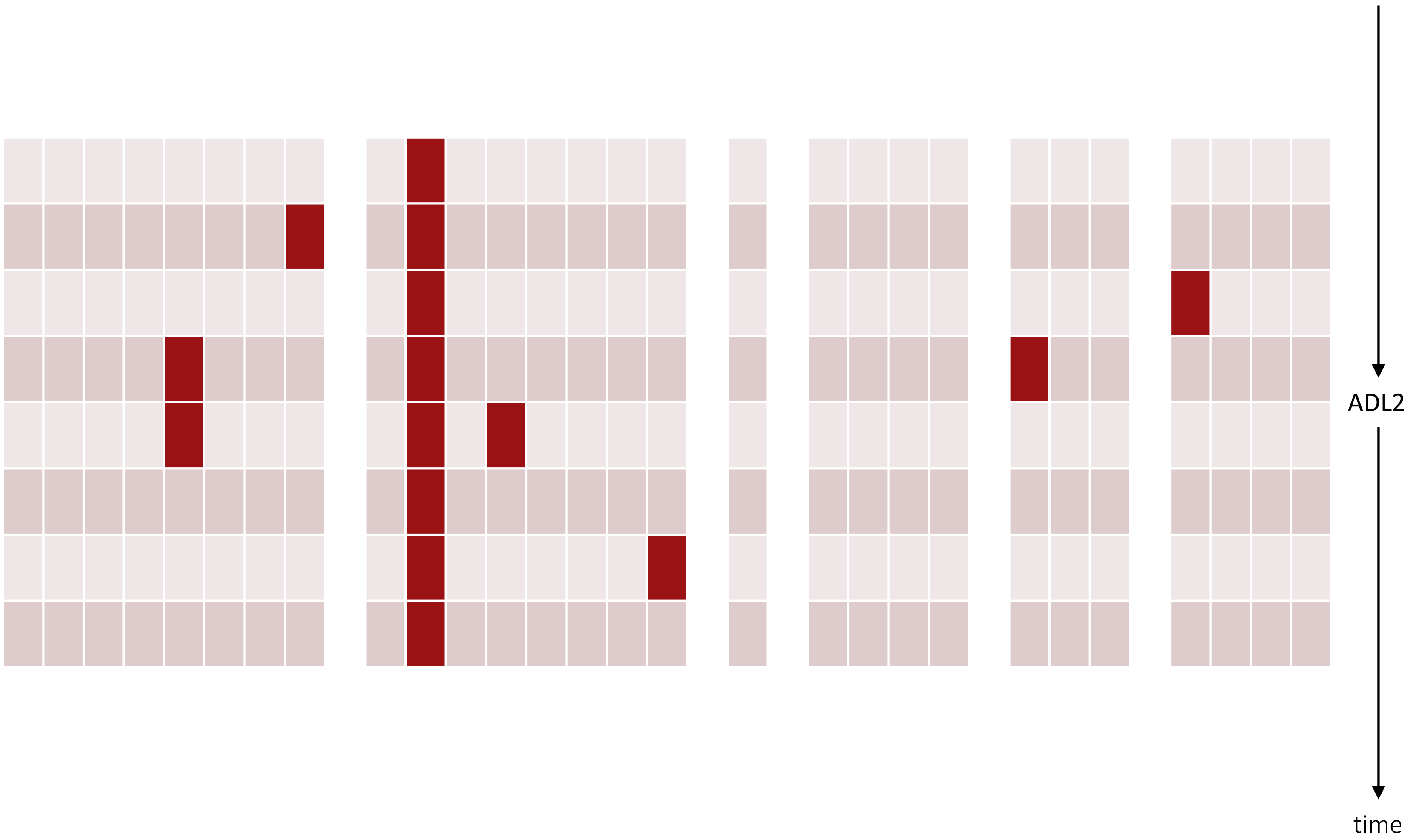
Preprocessing

Discard of useless columns, keeping on-body sensor signals



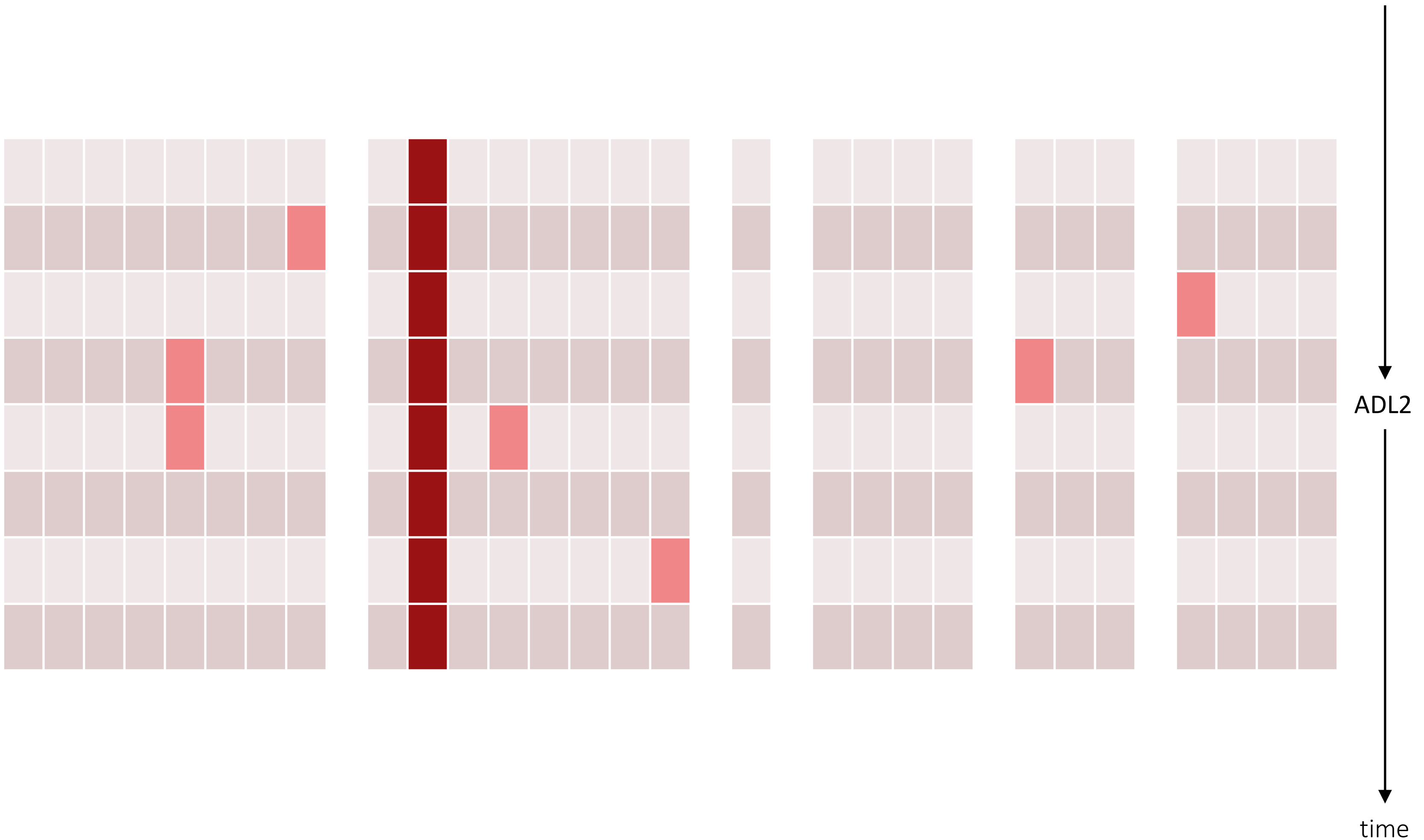
Preprocessing

Cutting of initial and final NaNs



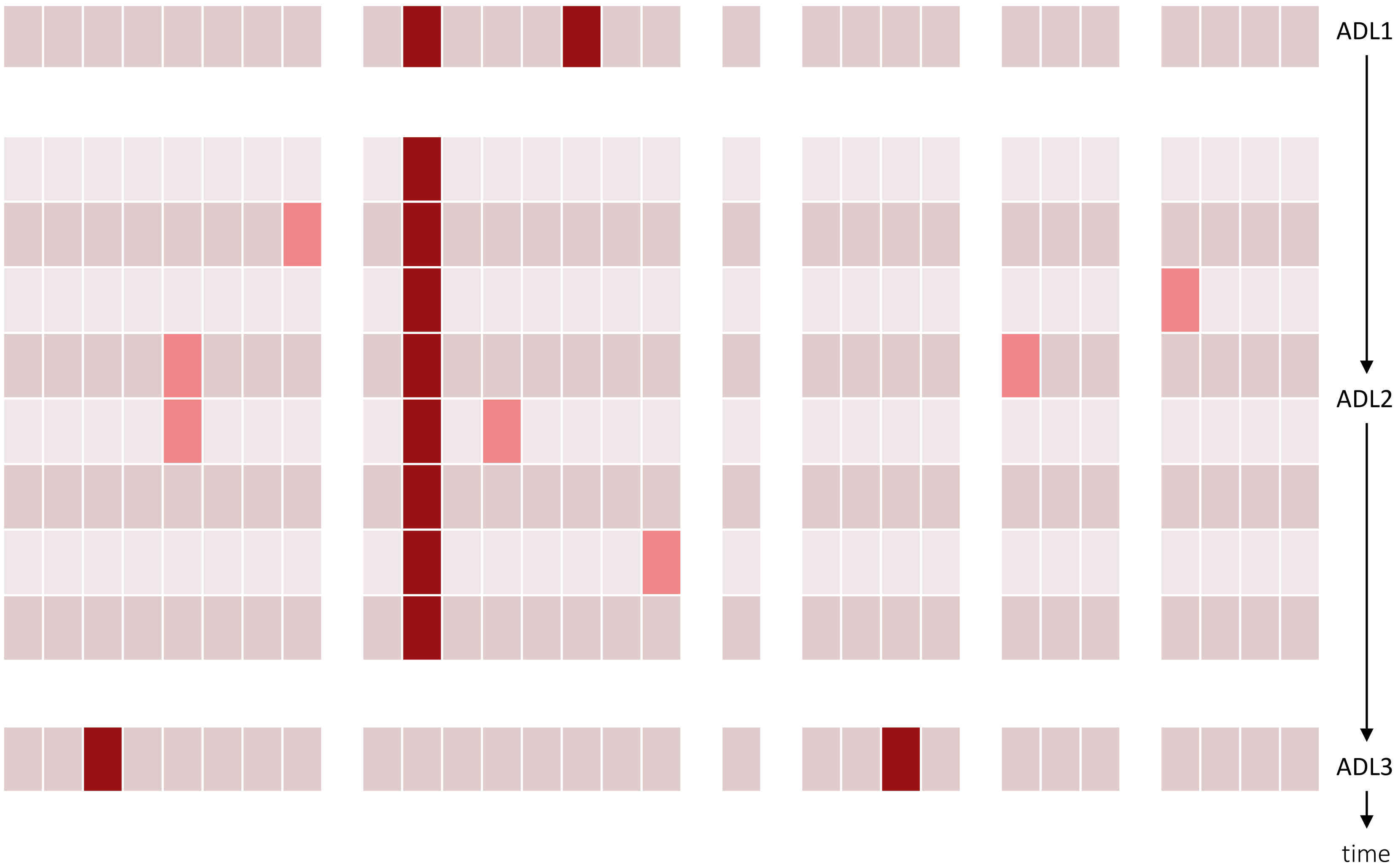
Preprocessing

Interpolation



Preprocessing

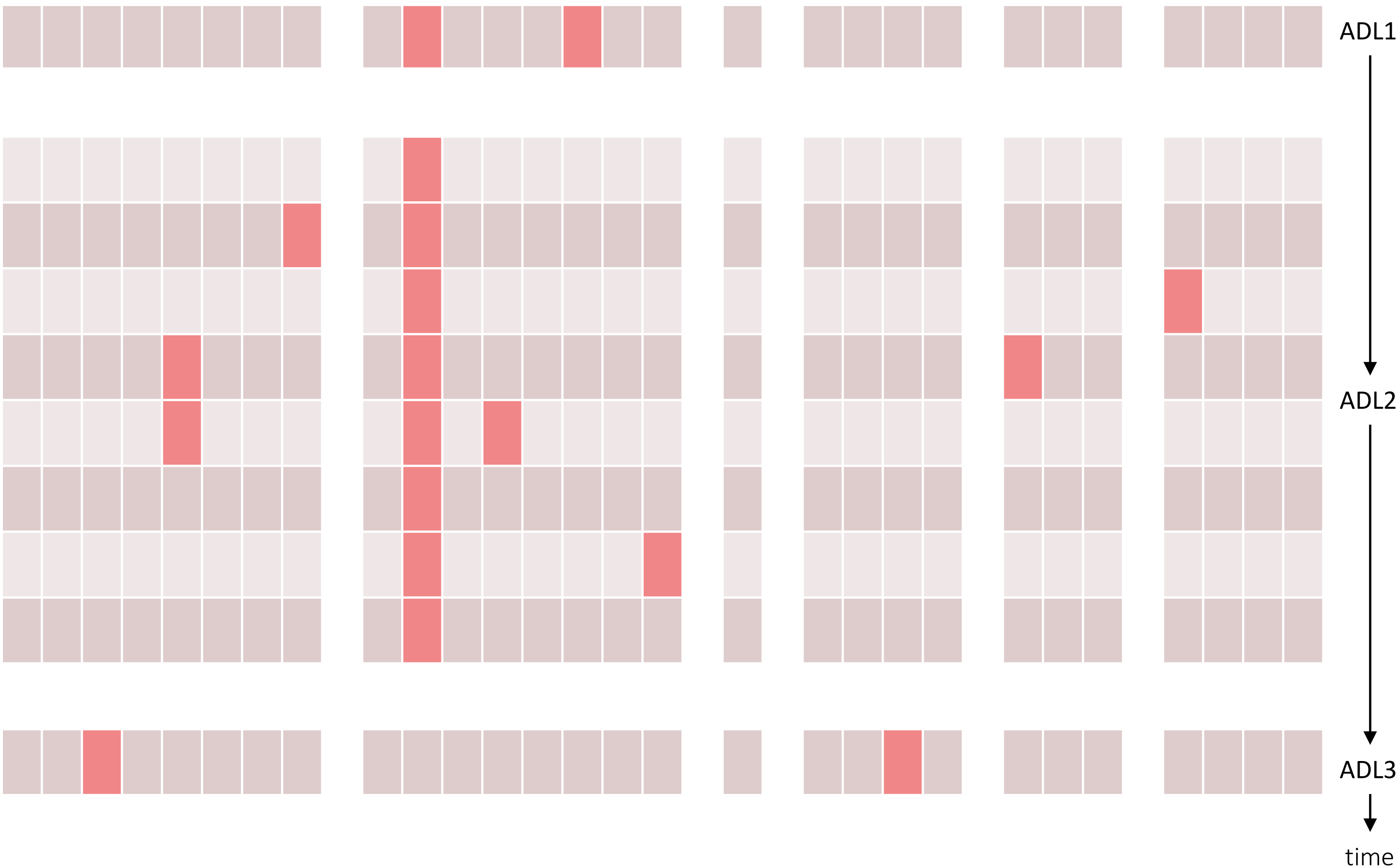
Concatenation



HUMAN DATA ANALYTICS

Preprocessing

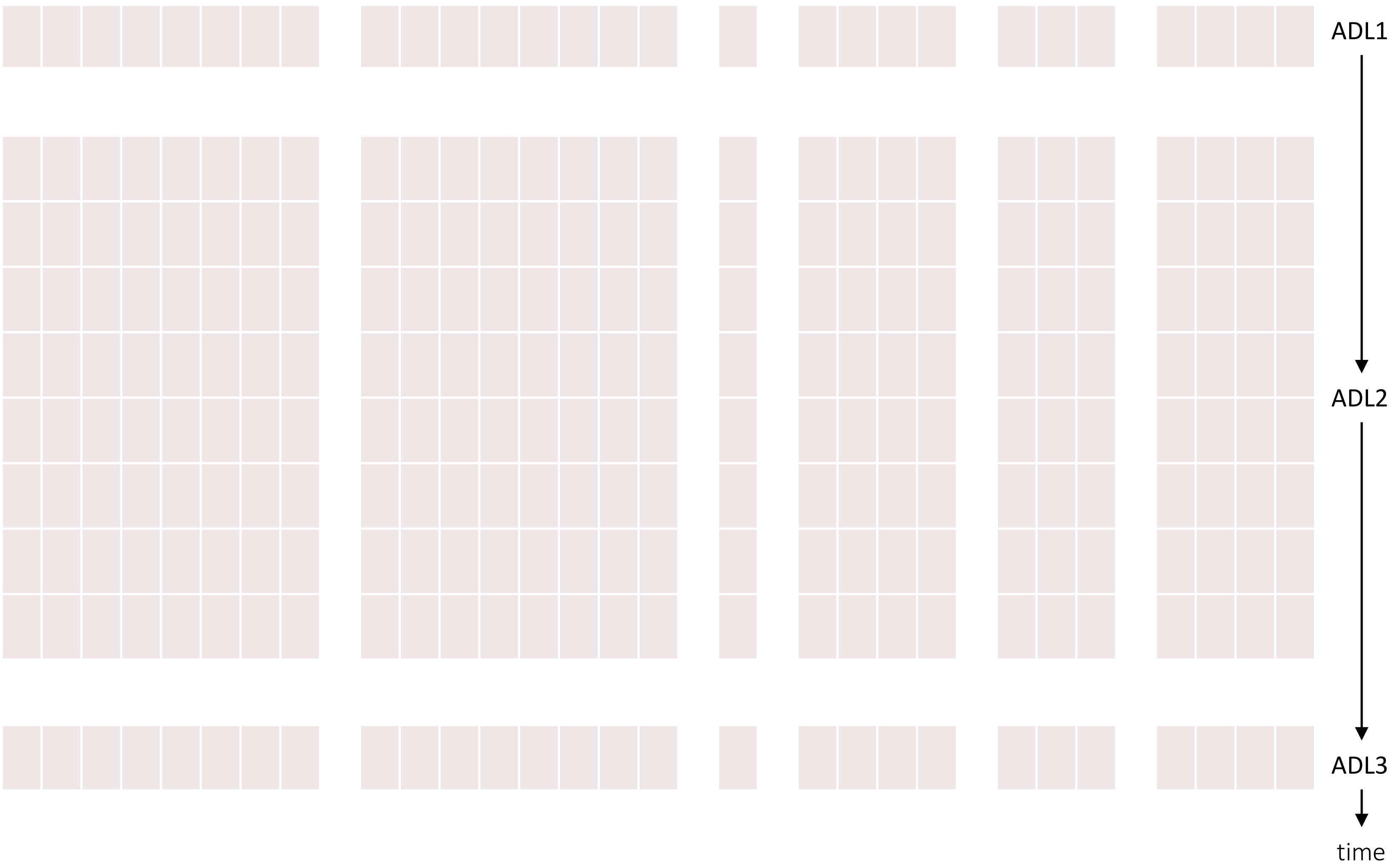
Setting Nan columns to zero



HUMAN DATA ANALYTICS

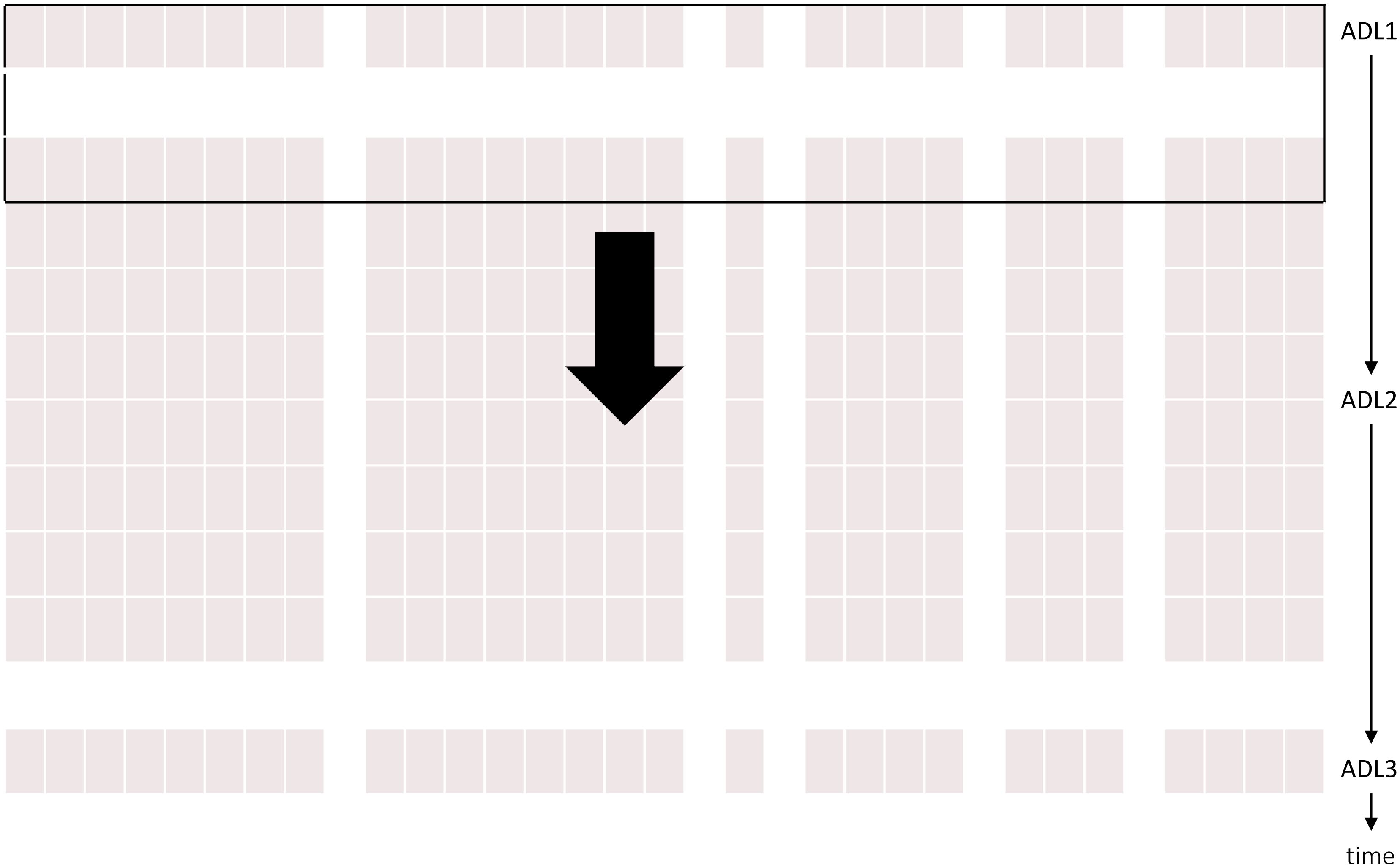
Preprocessing

Normalization



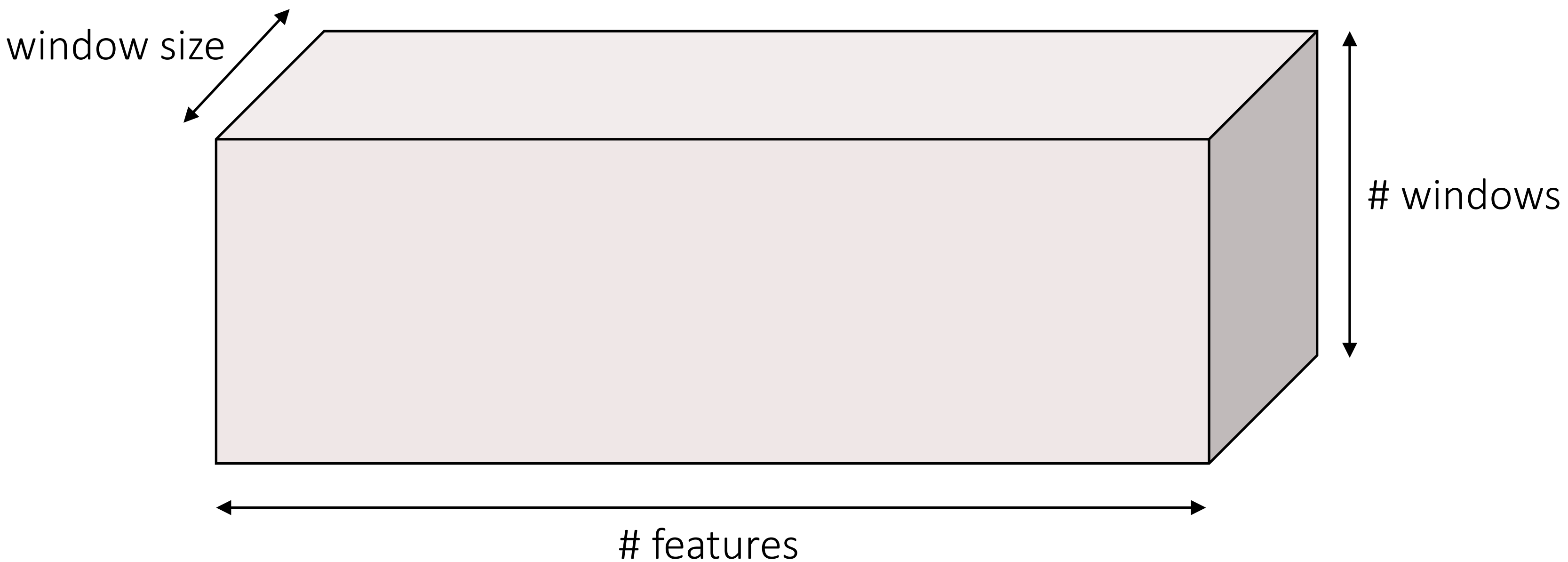
Preprocessing

Shaping (windowing)



Preprocessing

Shaping (windowing)



Training set	ADL1	Test set	ADL4
	ADL2		ADL5
	ADL3		
	Drill		

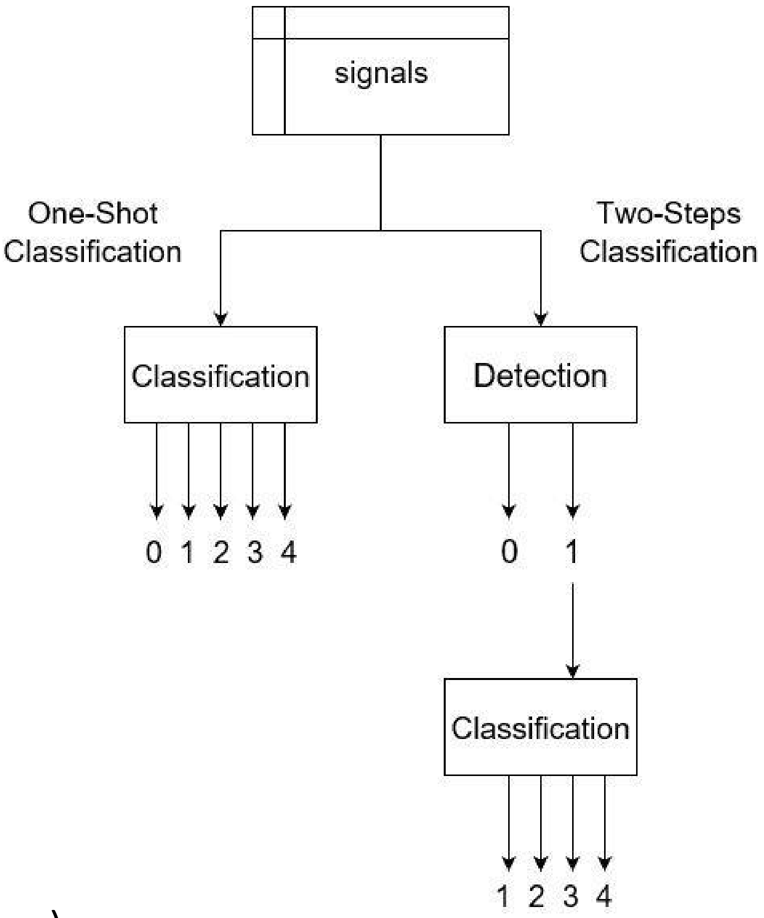
Models

Layer type \ Model name	Convolutional	Recurrent (LSTM)	Fully connected
Conv	3	0	2
Conv1DRec	1	2	2
Conv2DRec	1	2	2
ConvDeepRec	3	2	2

In our code:

Conv → Convolutional
Conv1DRec → Convolutional1DRecurrent
Conv2DRec → Convolutional2DRecurrent
ConvDeepRec → ConvolutionalDeepRecurrent

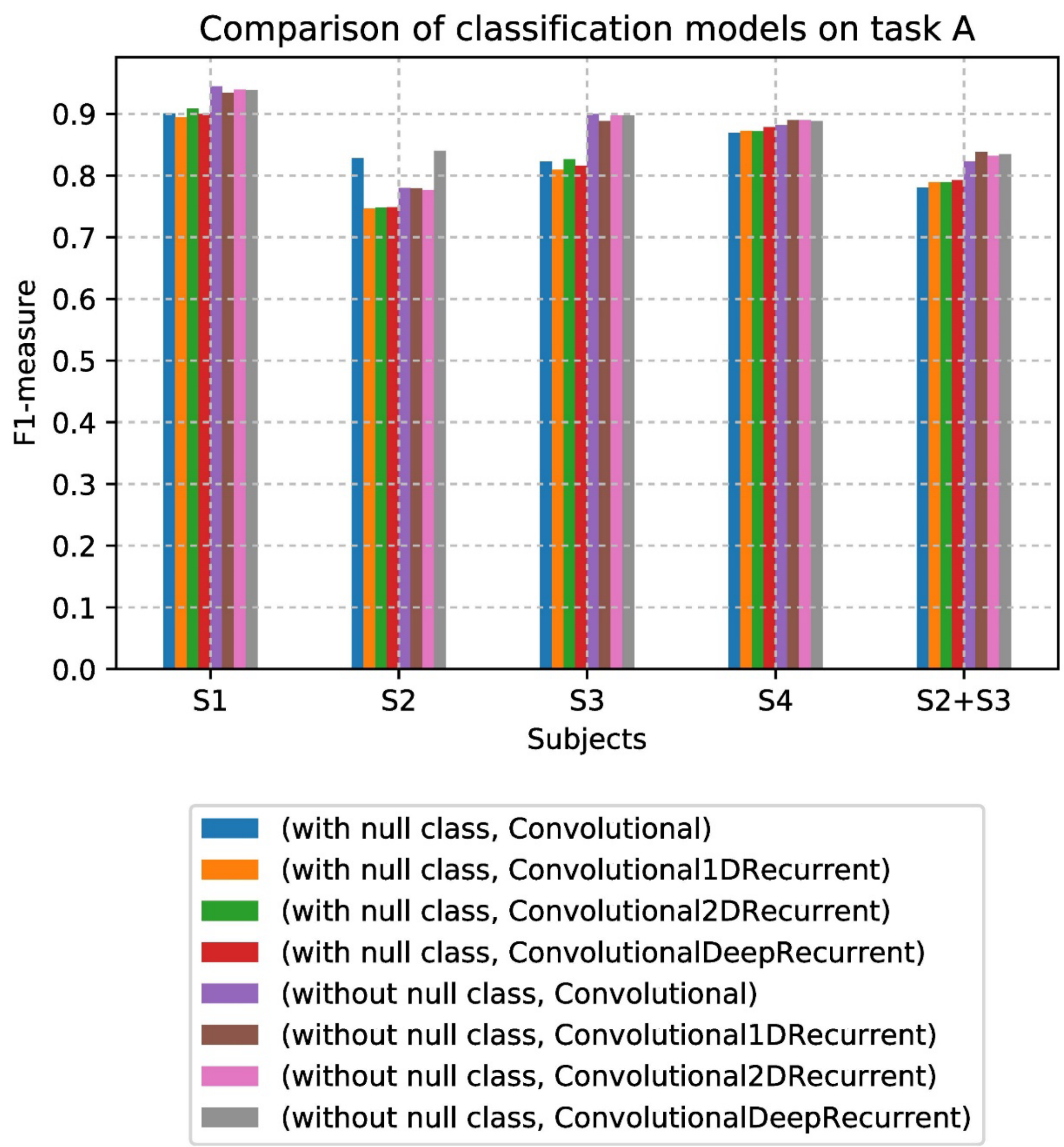
Pipelines



3 types of classification for each model:

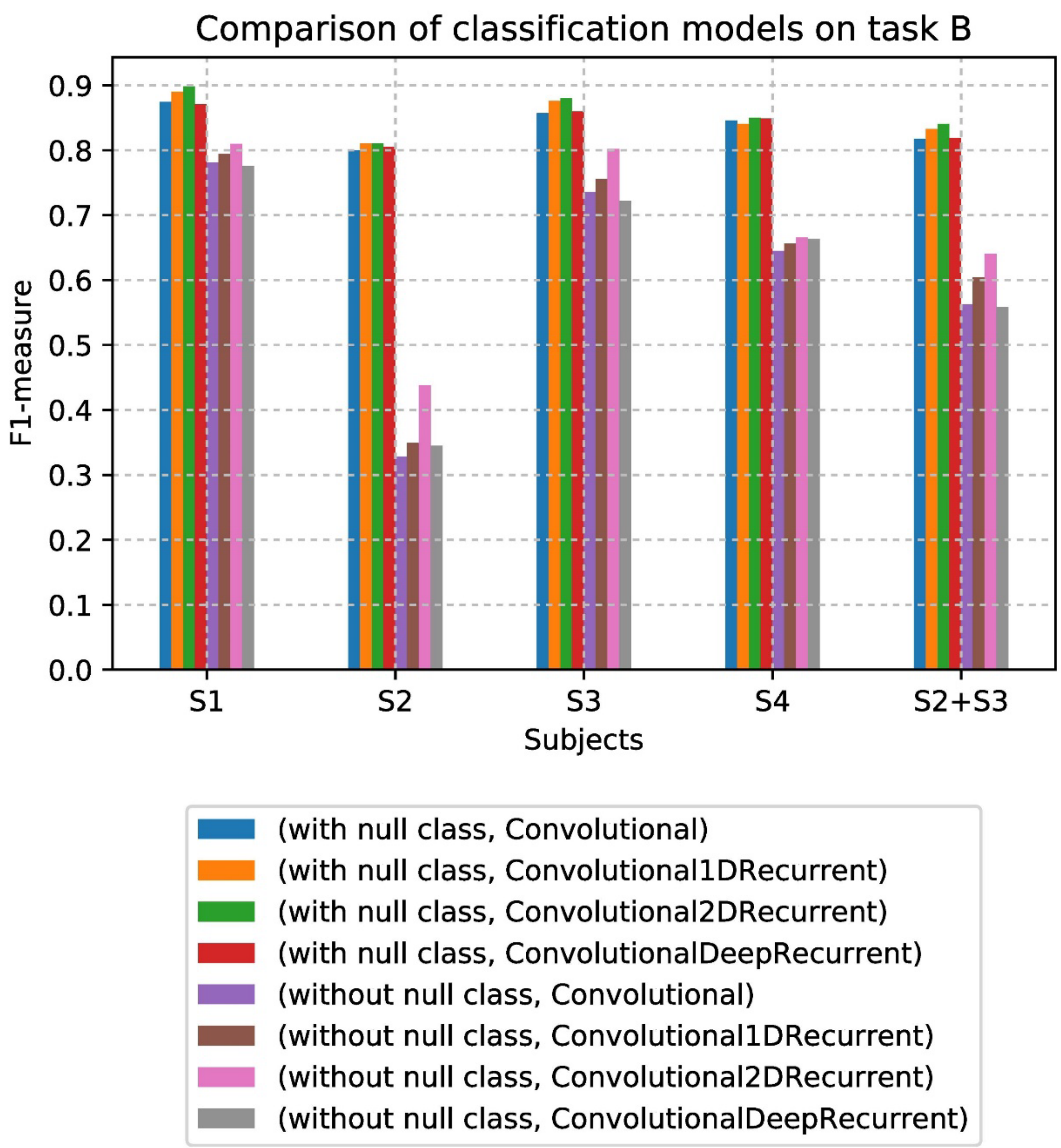
- One-Shot
  - Classification (n+1 class)
- Two-Steps
  - Activity detection (binary)
  - Activity classification (n-class)

Results on locomotion



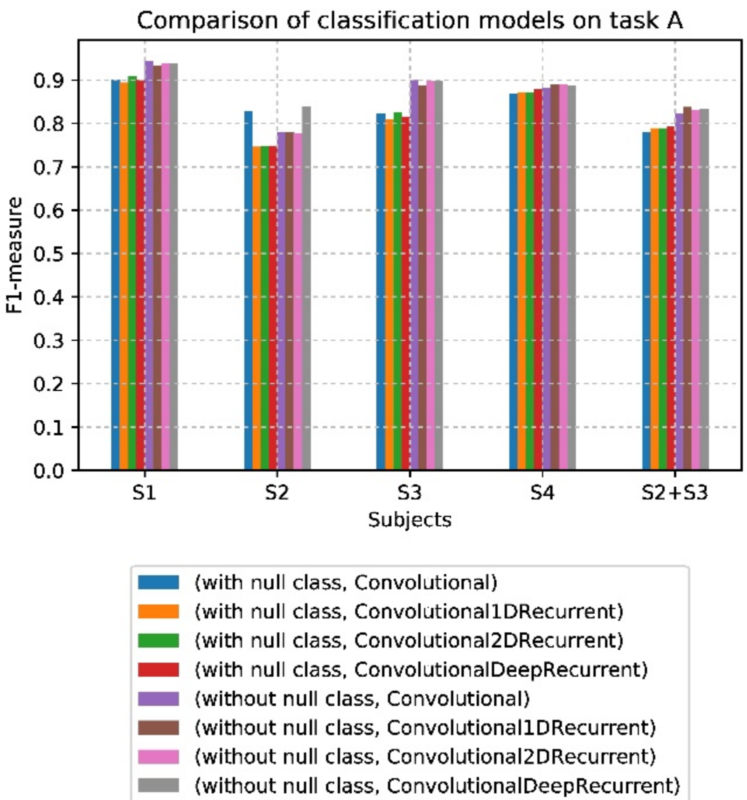
HUMAN DATA ANALYTICS

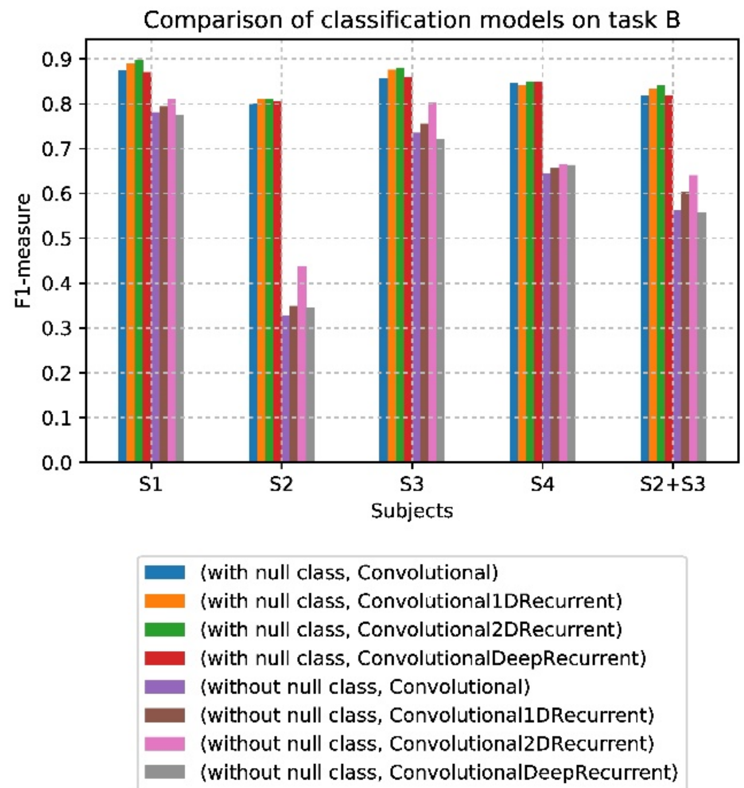
Results on gestures



HUMAN DATA ANALYTICS

Conclusions





- No clear best choice
- Class imbalance
- Cascade implementation … notebook

These problems could be addressed in future work

HUMAN DATA ANALYTICS

Deep Learning Techniques for Gesture Recognition: Dealing With Inactivity

Matteo Drago, Riccardo Lincetto
Dept. of Information Engineering, University of Padova, Italy

Prof. Michele Rossi

matteo.drago@studenti.unipd.it
riccardo.lincetto@studenti.unipd.it

September, 5, 2018