The Art of
Sport Activity Classification

Presenter
Matarmaa Jarno Olavi
Master of Artificial Intelligence

Literature overview

- Human Activity Recognition (HAR) is well studied area in recent years, achieving average accuracies of **92** % (*Demrozi et al., 2020*).
- Traditional HAR studies are conducted using inertial sensor data (*basically, a set of 3-dimensional (x,y,z) accelerometer sensors*)
- Amount of activity types varies from 5 to 20, containing *sitting down*, *standing up*, *laying down*, etc., among most common sports.
- Classification tasks have been conducted using wide variety of CML and Neural Network models

Study objectives

- Providing alternative method for SAC
- Classifying sport activities from the extracted sensor features when original sensor data is not available
- Investigating the effect on accuracy when using different data structure and data types
- Reach a competitive level of accuracy with previous HAR studies

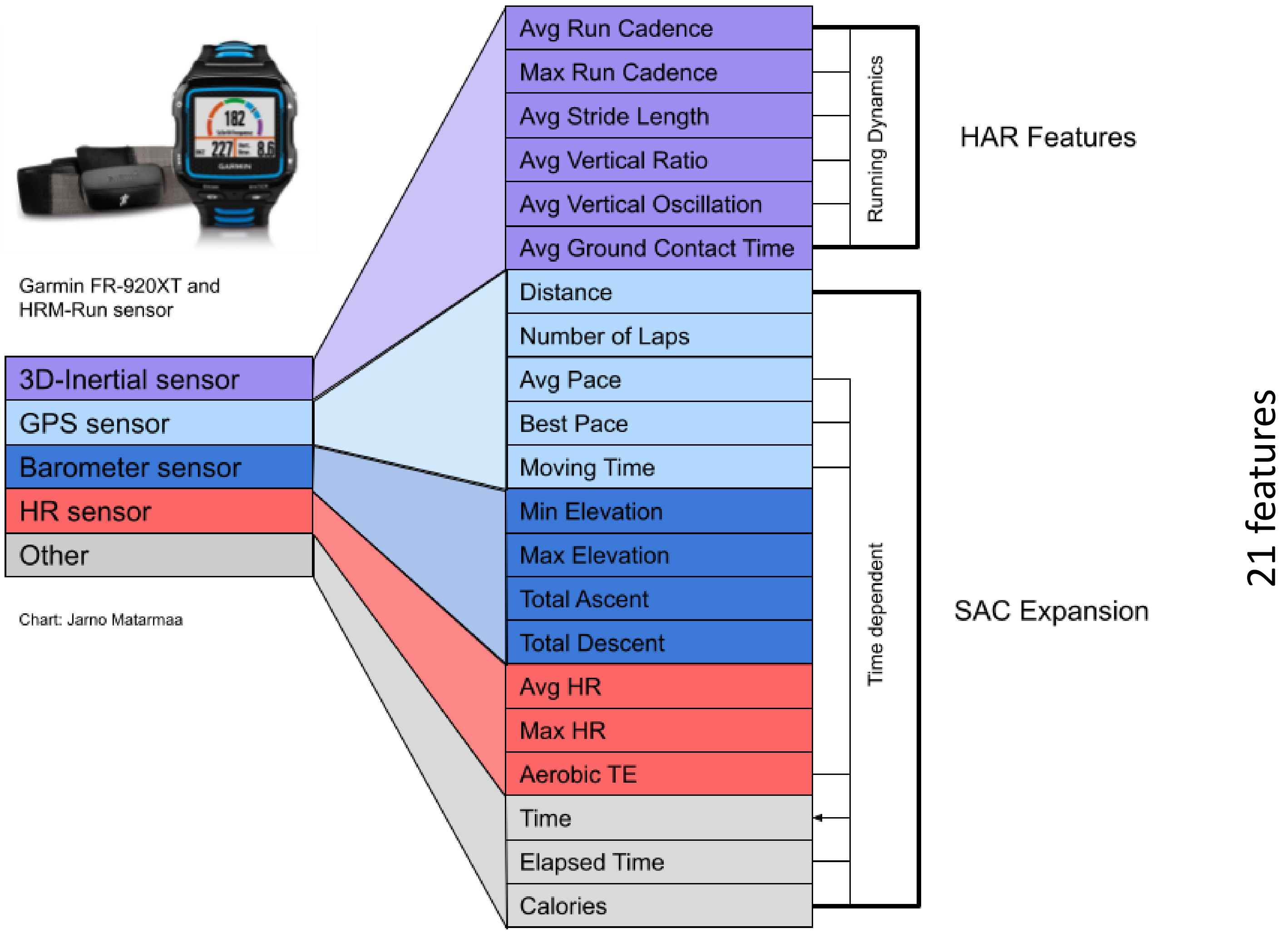
The relevance of the study

- The daily usage of activity trackers, sport watches, and other devices to record daily and sport activities has increased tremendously in past ten years
- The main reasons are the development of battery technology and small size hardware with a low energy consumption such as LED displays
- Human activity recognition has become a common function and is enabled by several device manufacturers

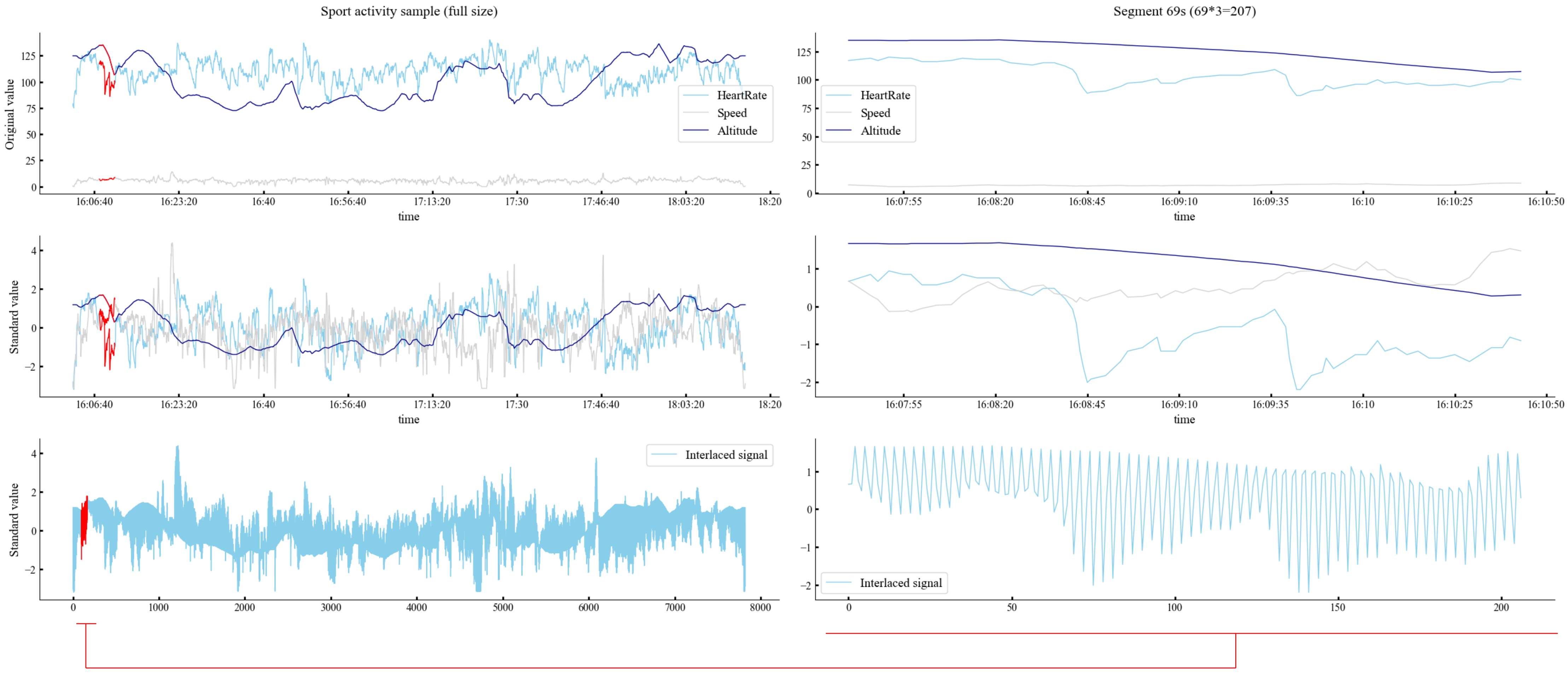
Materials and methods

- Two type of datasets from a single athlete
  - **Training history dataset** for S-CML with extracted features **(21)**
  - 3-dimensional pure **time series signal dataset** for TSC
  - Original dataset has **297 activities** in **5 diverging sport categories**
  - Basically, the very same data in two different structure
  - Classes/categories are {**biking, running, walking, skiing, roller-skiing**}
- Adopted classifiers
  - **9** Standard CML classifiers from *Scikit-learn* (sklearn)
  - **11** Time Series Classifiers from *Sktime API*
- Evaluation metrics
  - Accuracy (Mean), Model training time
  - Precision, Recall, F1-score
  - ROC-AUC, Stability (accuracy variance)

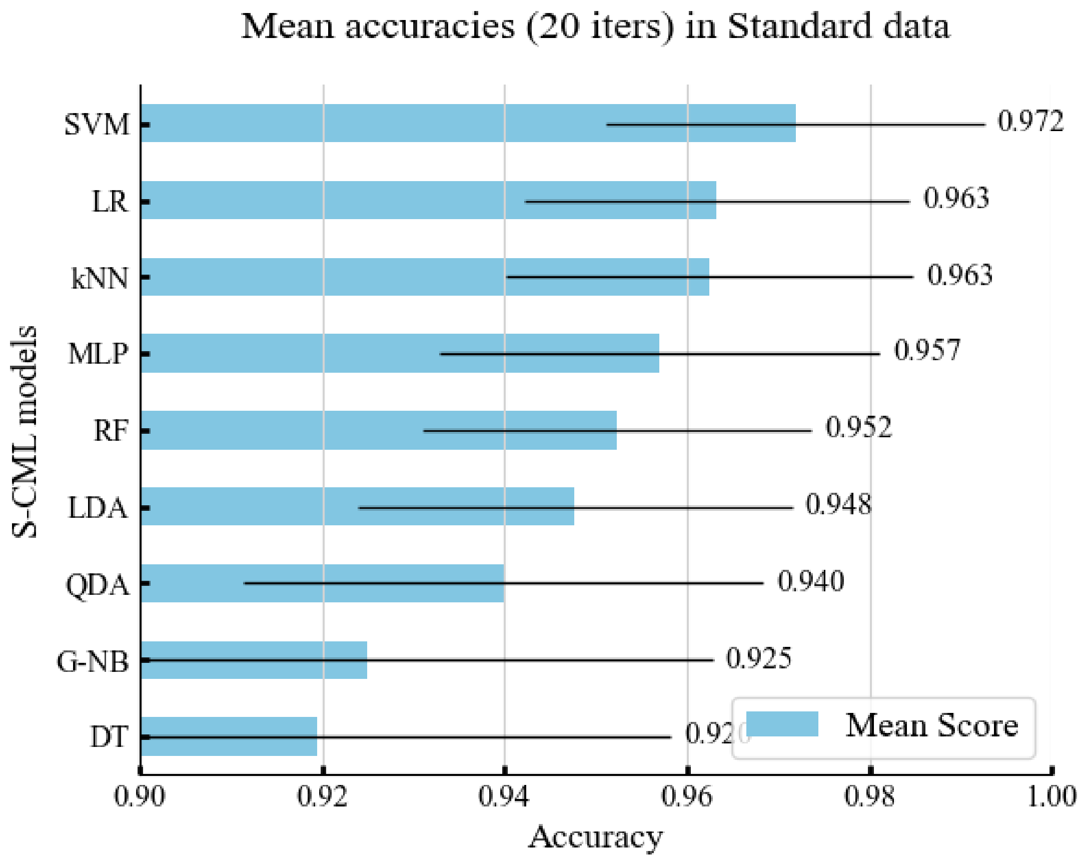
S-CML Dataset Features

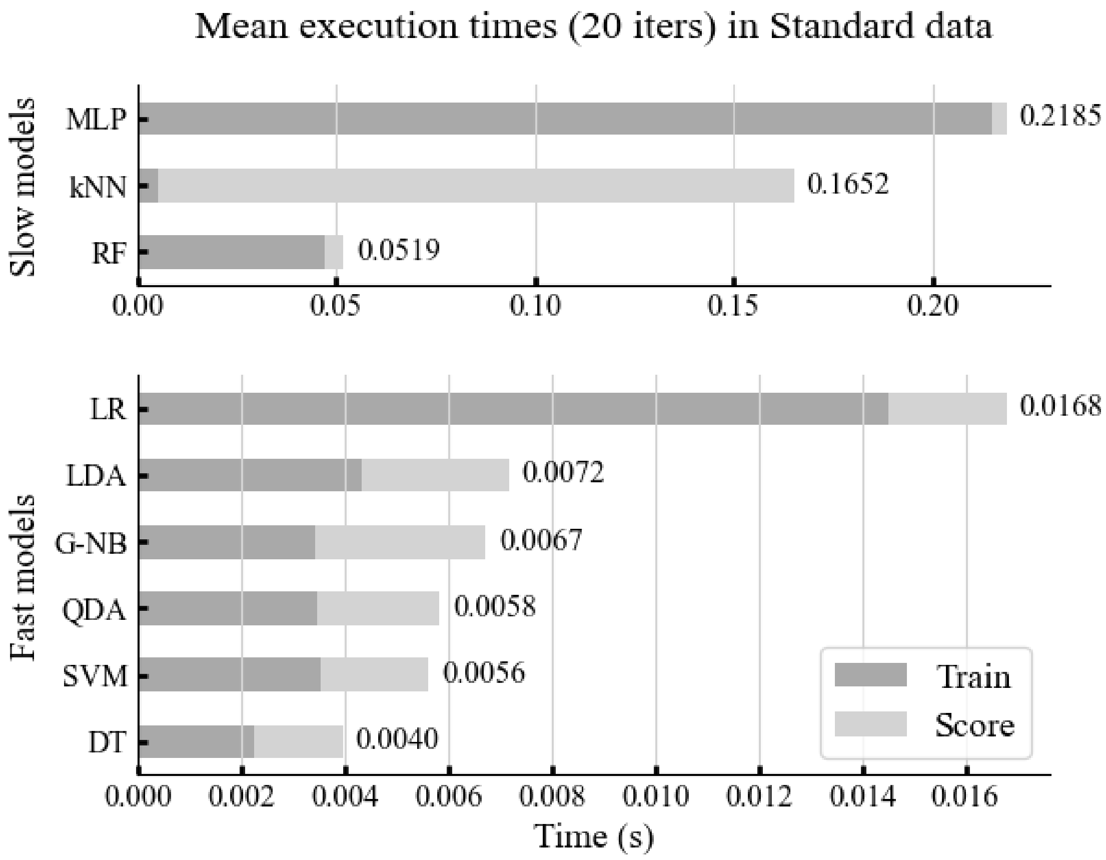


Interlacing method of multivariate signals (TSC)

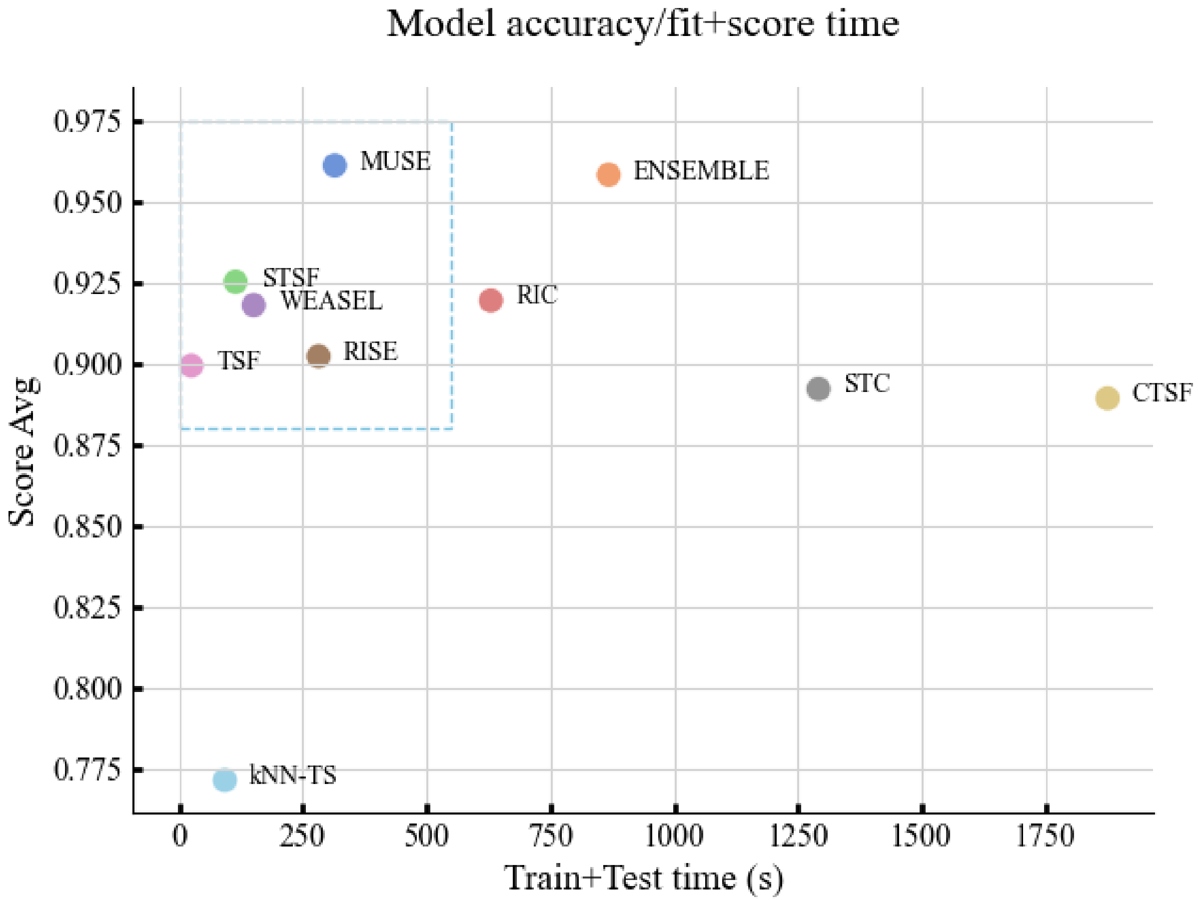


Results: S-CML





Results: TSC

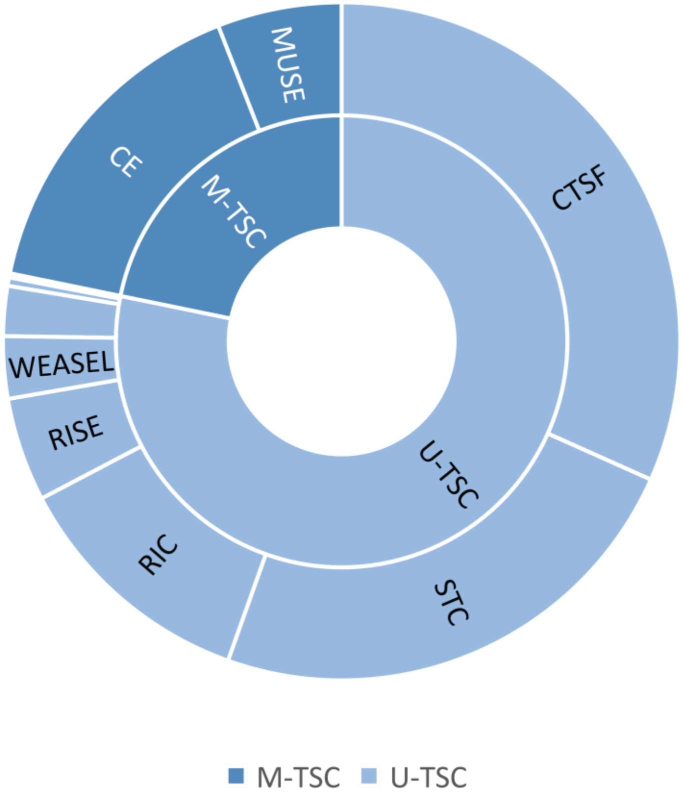


Performance map Accuracy-Time

Results: Compilation table

Classifier	Type	ROC-AUC	Best Score	Score Variance	Mean Score	RTC [0,10]
MUSE	M-TSC	0.994	**0.966**	0.004	0.961	3
ENSEMBLE	M-TSC	0.995	**0.961**	0.002	0.958	8
STSF	U-TSC	0.990	**0.931**	0.004	0.925	1
RIC	U-TSC	0.985	**0.927**	0.006	0.920	6
WEASEL	U-TSC	0.981	**0.918**	0.000	0.918	1
GB	S-CML	-	**0.918**	0.004	0.918	0
MLP	NN	-	**0.918**	0.000	0.912	0
RISE	U-TSC	0.980	**0.905**	0.002	0.902	2
RF	S-CML	0.980	**0.905**	0.004	0.902	0
TSF	U-TSC	0.979	**0.901**	0.002	0.899	0.2
STC	U-TSC	0.973	**0.897**	0.004	0.892	10<
CTSF	U-TSC	0.978	**0.892**	0.002	0.889	10<
SVM	S-CML	-	**0.884**	0.000	0.884	0
DT	S-CML	0.898	**0.875**	0.013	0.871	0
LR	S-CML	-	**0.853**	0.000	0.853	0
kNN	S-CML	-	**0.849**	0.000	0.849	0
LDA	S-CML	-	**0.789**	0.000	0.789	0
G-NB	S-CML	-	**0.784**	0.000	0.784	0
kNN-TS	U-TCS	0.829	**0.772**	0.000	0.772	<1
QDA	S-CML	-	**0.573**	0.000	0.573	0

Relative Time Complexity (RTC)



Conclusions

- Three diverging methods (S-CML, U-TSC, and M-TSC) were successfully applied for *a retrospective personalized supervised sport activity classification* of a single athlete
- The observed results are mainly good, up to 96,6%. HAR average (92%) (*Demrozi et al., 2020*)
- Model selection depends on the available data and application objectives
  - **MUSE** (96,6%) can be gently recommended for **the multivariate time series data**
  - In certain circumstances, **TSF** as a **fast model** could be more preferable than **MUSE**, and then solving this problem in **univariate space** can take place
  - **S-CML** as a very fast and effective method with low computation requirement could be suggested **always when the appropriate dataset is available,** but it might be prone to interpersonal differences, and therefore results not generalizable

Right Questions

1. What’s the significance of constructing **interlaced signals** instead of using simple **column concatenation** method?
2. In which cases pure **time series signals** could be especially useful compared to **extracted features**?
3. Why we should classify sports from **heart rate, speed,** and **altitude** features instead of using **inertial sensor data** which have produced a great results in previous studies?
4. Are the results **generalizable** and can we expect the **same accuracy level** among athletes with different physical and personal characteristics such as **gender, age, weight, length**, etc.?
5. Why we should have a specialized classification method for outdoor sports?
6. Why we should classify sports at all?
7. What makes it preferable to conduct a personalized sport classification?

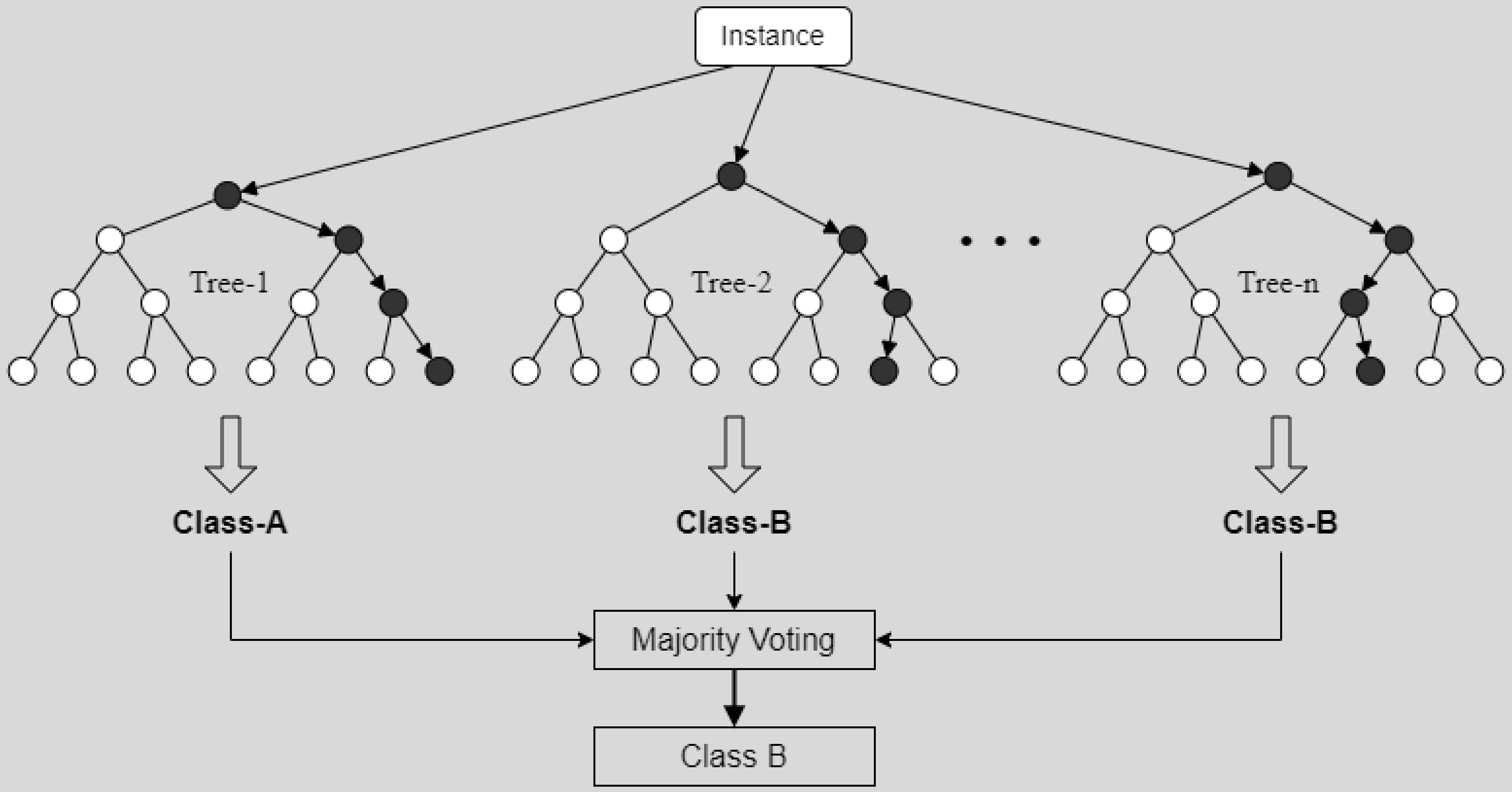
Discussion...

Appendices

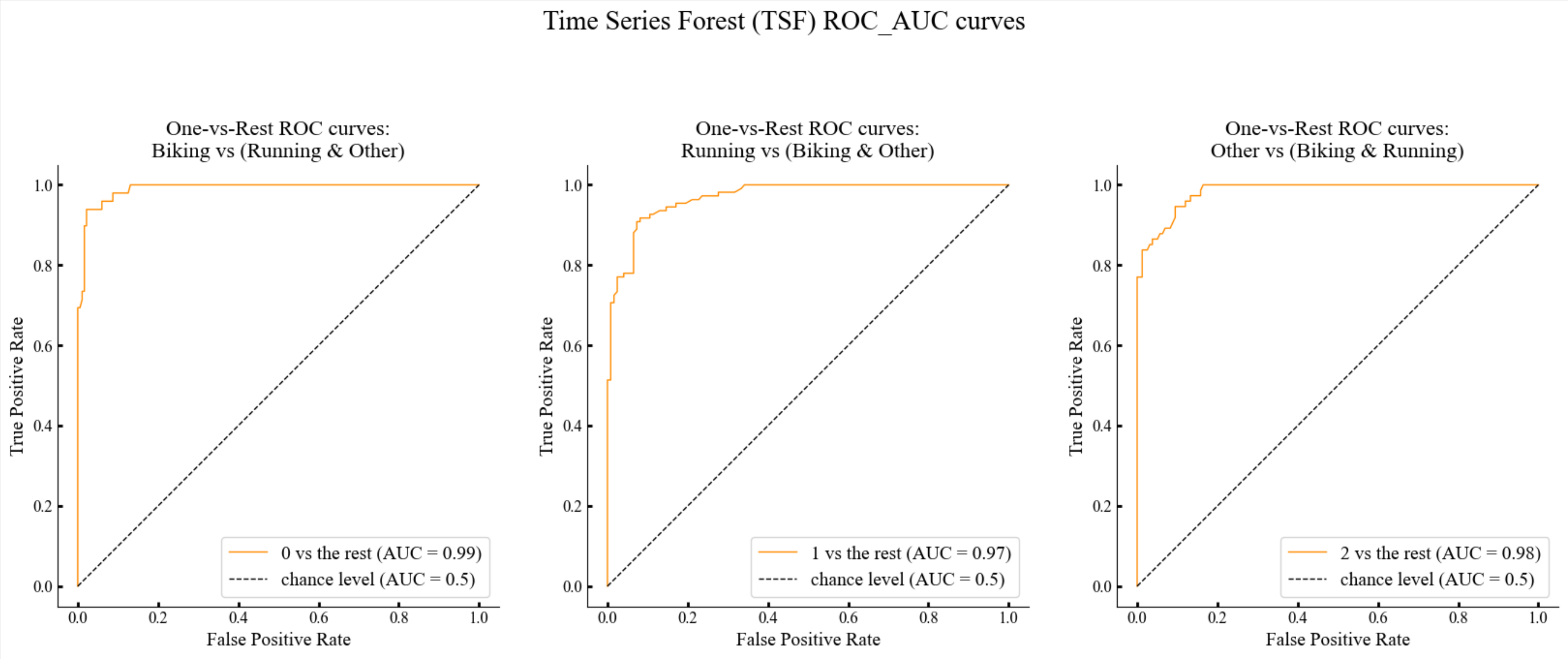
Why personalized SAC with a dataset recorder in uncontrolled non-laboratory environment

- Foerster et al. (1999) demonstrated accuracy drop for ambulation activities from 95.6% of a controlled data collection experiment to 66% of uncontrolled non-laboratory natural environment.
- Berchtold et al. (2010) identified an open debate on the design of any activity recognition model. Since according to some authors, people perform activities in a different manner as they differ on age, gender, weight, and so on, a specific recognition model should be built for each individual.

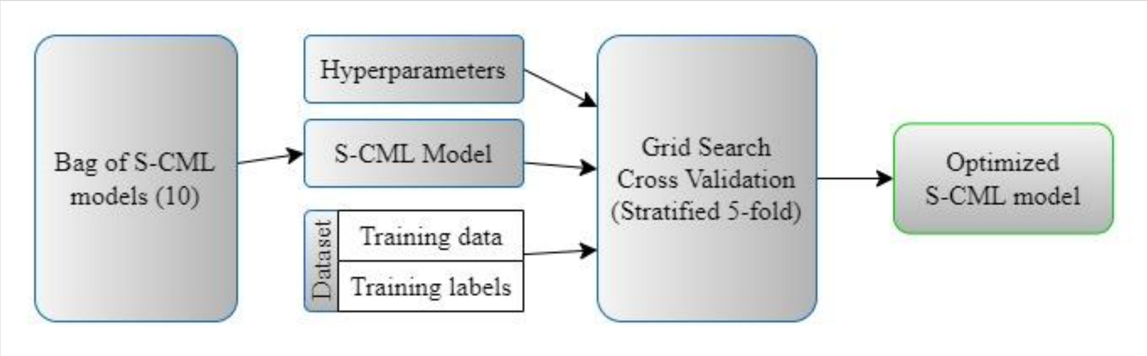
TSC - Random Forest



TSC - Classification quality metric ROC-AUC

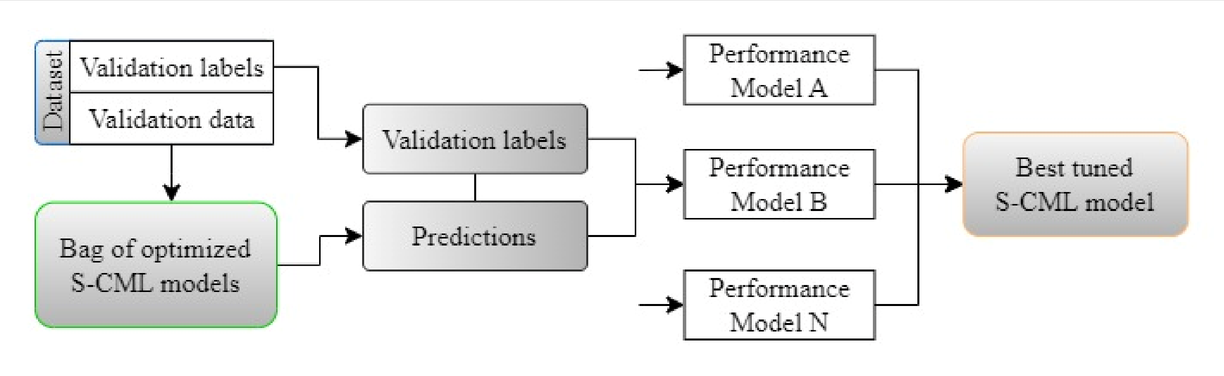


S-CML Pipeline



Step 1: Hyperparameter optimization

Step 2: Model selection



Category divergency in S-CML data