Уральский федеральный университет имени первого Президента России Б.Н.Ельцина
Институт радиоэлектроники и информационных технологий — РТФ

Уральский федеральный университет имени первого Президента России Б.Н.Ельцина

The Art of Sport Activity Classification

Докладчик
Matarmaa Jarno Olavi
Master of Artificial Intelligence

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 enabled by several device manufacturers

Current level and methods of problem solving

- 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*)
- Studies have been conducted in a controlled laboratory conditions
- Amount of activity types varies from 5 to 20, containing sitting, standing up, standing down, laying down, etc., among most common sports.
- Classification task have been conducted using wide variety of signal classification methods including Classical Machine Learning and Neural Networks

Purpose and problem description of the study

- Alternative ways to classify sport activities might be reasonable, for example, all the devices *do not have inertial sensors*, or such a *data is not available* for external analysis of third parties.
- The purpose of the study is to tackle false human sport activity labeling problem when inertial sensor data is not available
- Wrong sport labeling is problematic for a several reasons:
  - Misleading training guidance based on the training history
  - Distortion in general personal data statistics
  - Segment leaderboards loses their validity if one can use a bike to get a better place in a running segment

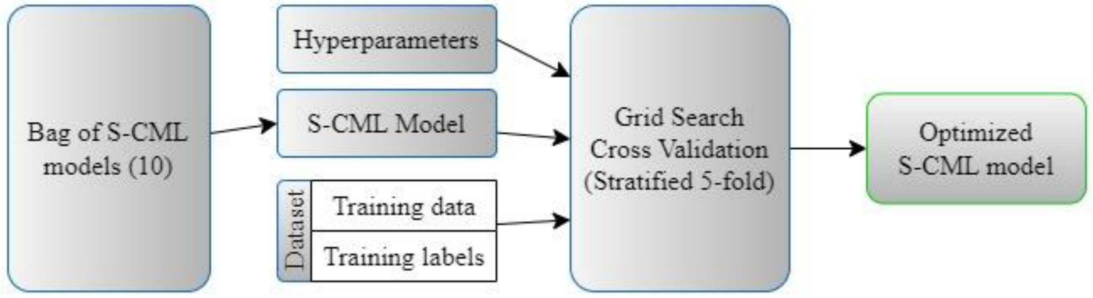
Materials and methods

- Two type of datasets from a single athlete
  - **Training history dataset** for S-CML with extracted features
  - 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

- Adopted classifiers
  - **9** Standard CML classifiers from *Scikit-learn* (sklearn)
  - **11** Time Series Classifiers from *Sktime API*

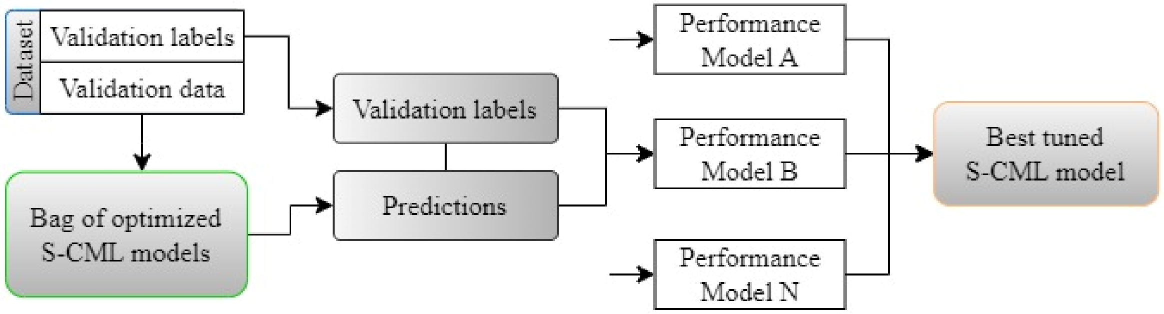
- Evaluation metrics
  - Accuracy (Mean), Model training time
  - Precision, Recall, F1-score
  - ROC-AUC, Stability (accuracy variance)

Materials and Methods: S-CML

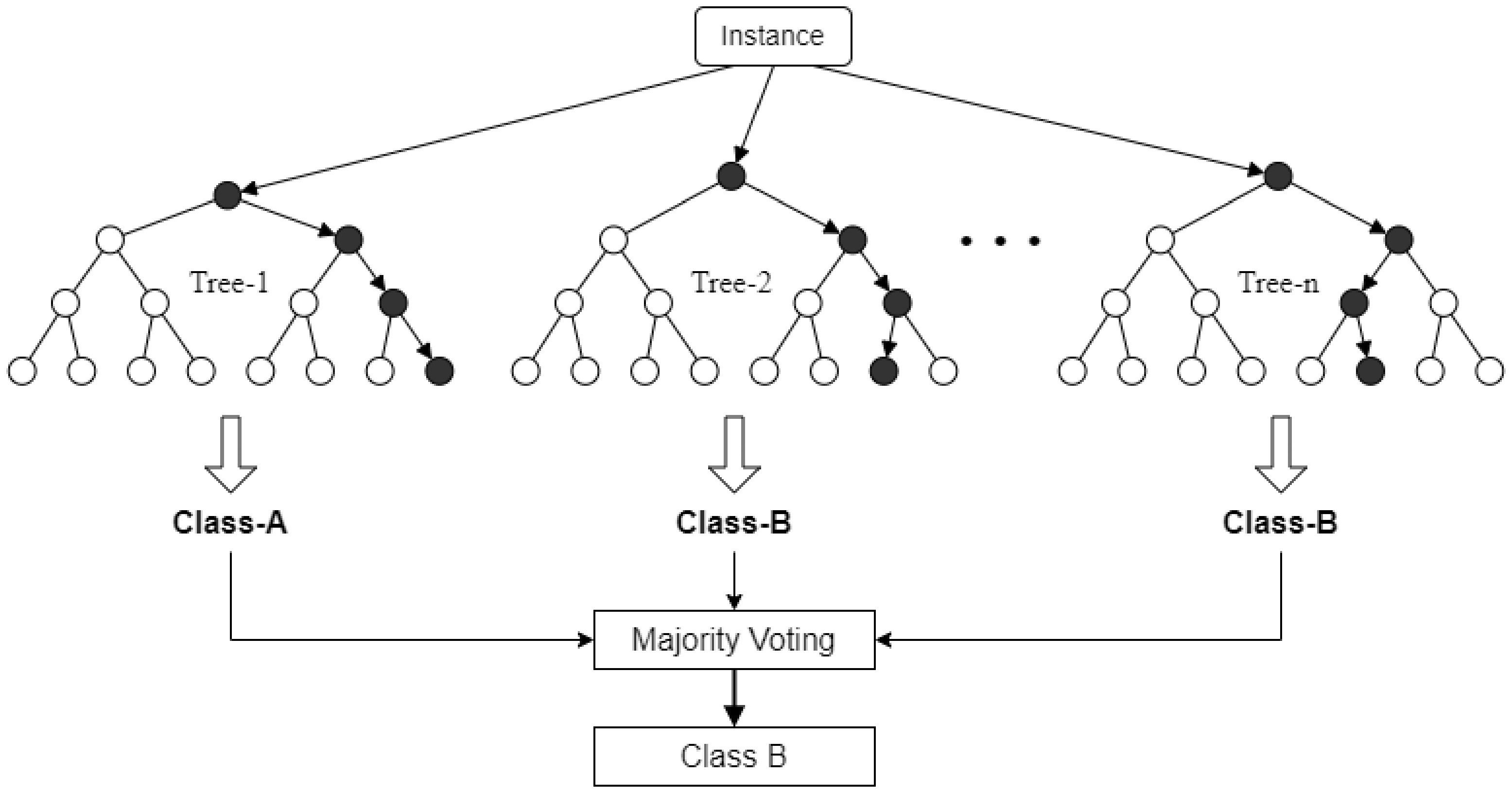


Step 1: Hyperparameter optimization

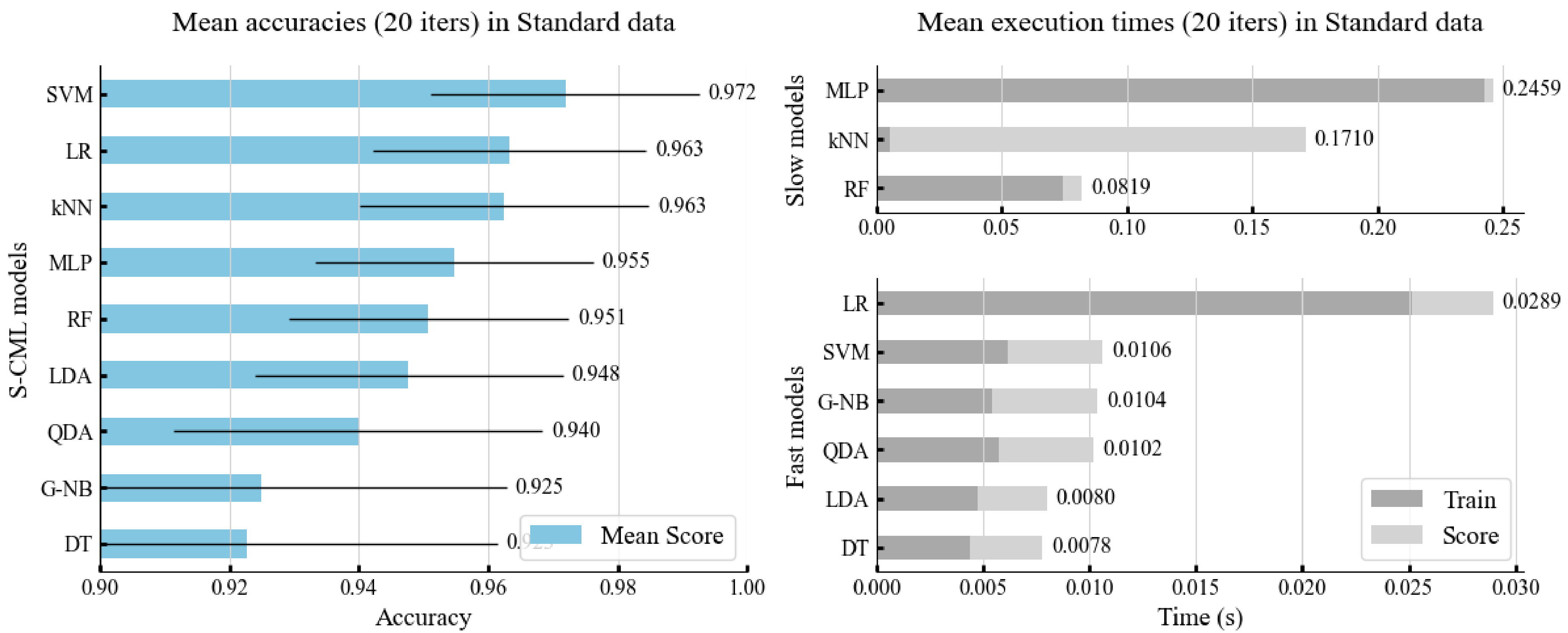
Step 2: Model selection



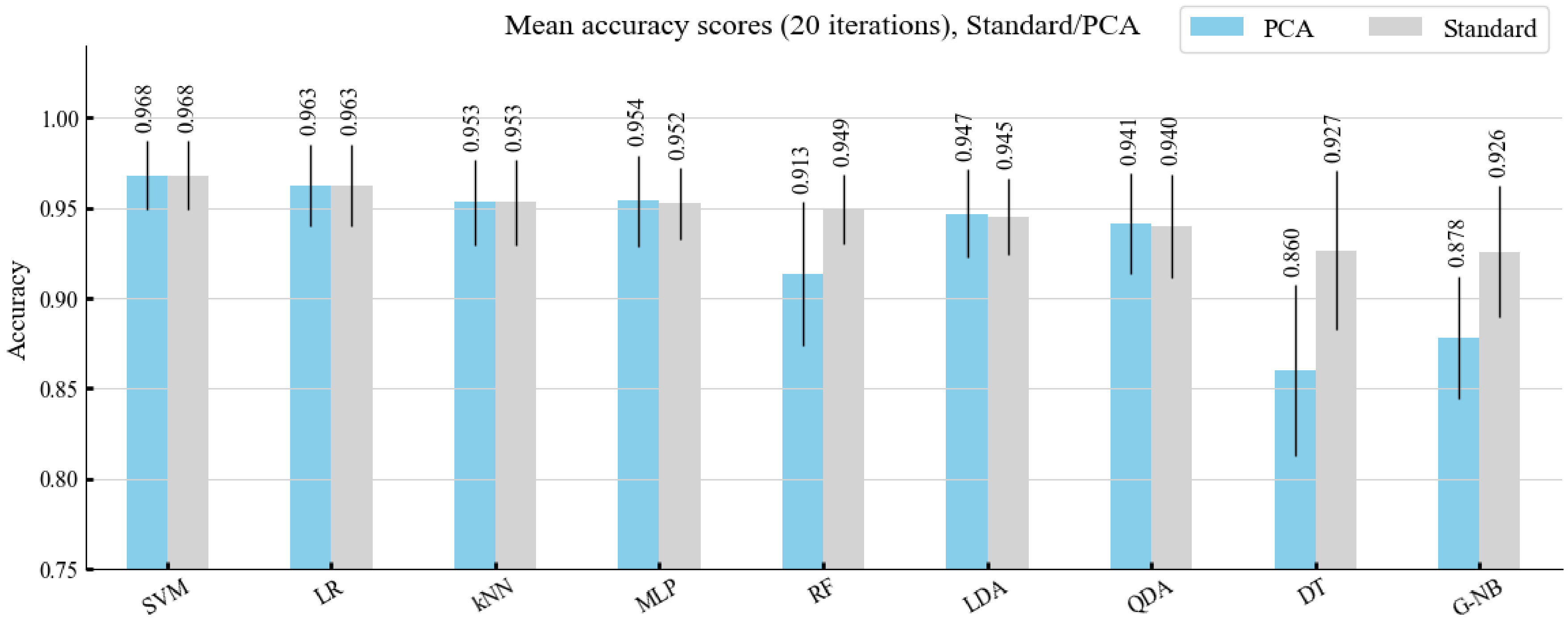
Materials and Methods: TSC – Random Forest



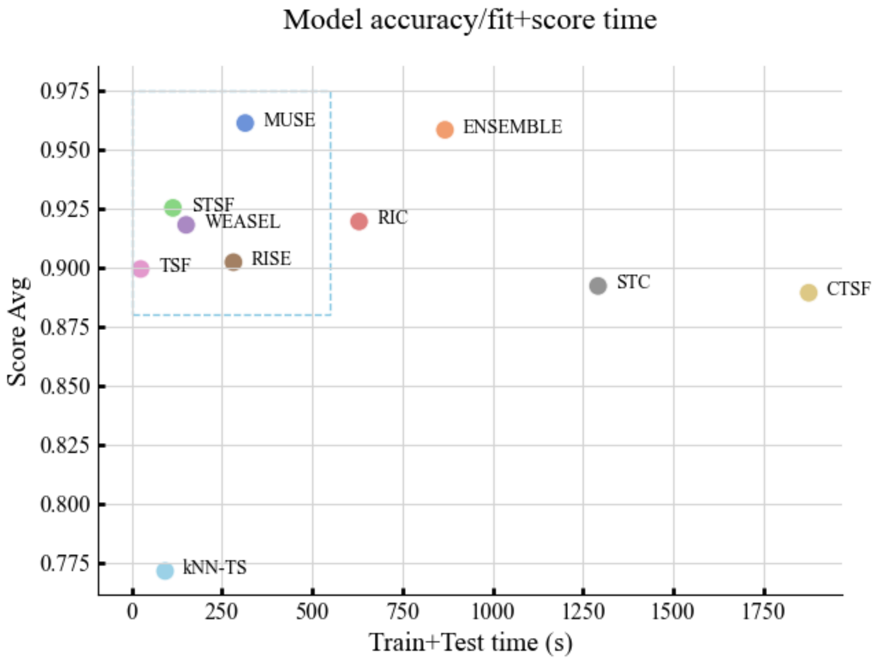
Results: S-CML



Results: S-CML



Results: TSC



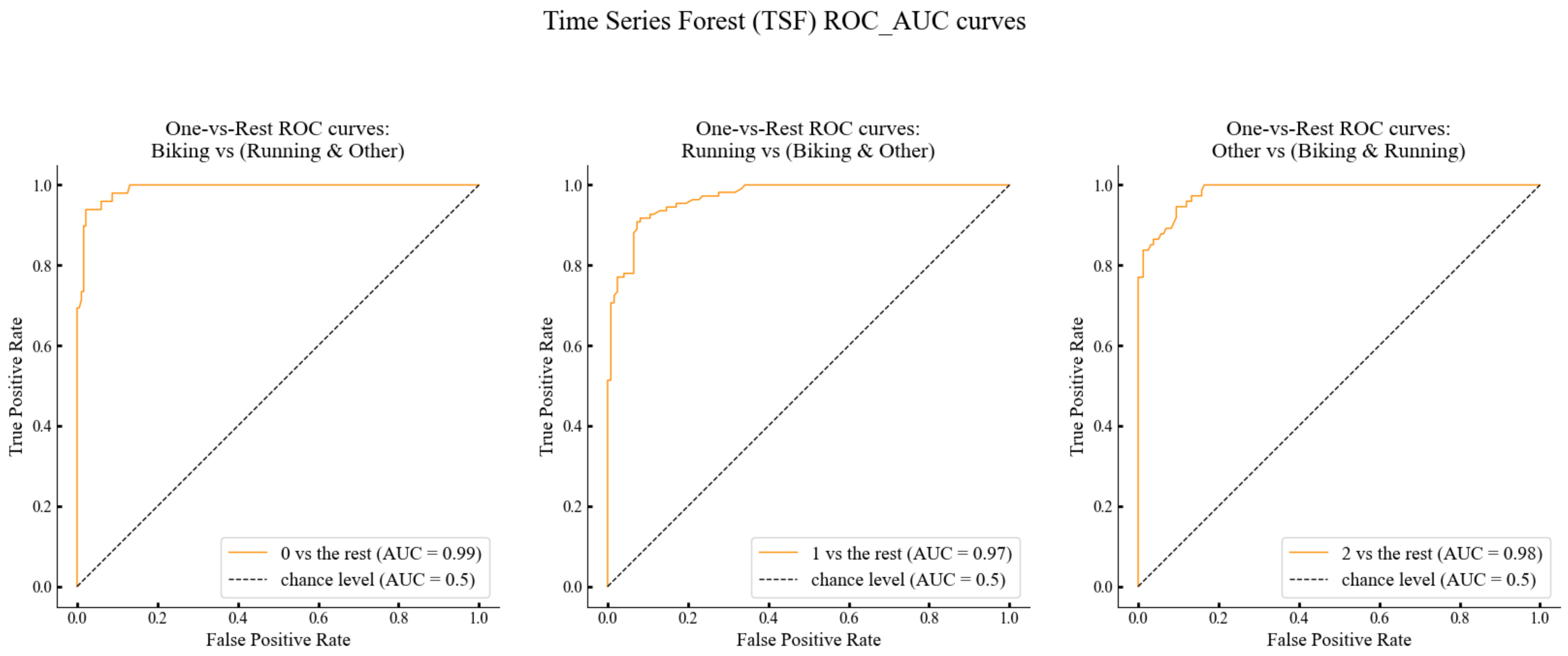
Performance map

Table 4.2K: Prediction results for eight most misclassified segments. Correct predictions are higlighted by grey color. First column indicates specific segment index in the test data (total of 232).

Test data index (232)	TSF	STSF	RISE	RIC	STC	kNN-TS	CTSF	WEASEL	Correct	Errors
220	1	1	1	1	1	1	1	1	**2**	**8**
62	0	0	0	0	0	0	0	0	**1**	**8**
71	0	0	0	0	0	0	0	0	**1**	**8**
97	1	1	1	1	1	1	1	0	**0**	**7**
17	1	0	1	1	1	1	2	1	**0**	**7**
208	1	1	1	1	1	2	1	0	**0**	**7**
138	1	1	0	2	1	1	1	2	**2**	**6**
195	1	1	1	1	1	2	1	2	**2**	**6**
Errors	**8**	**7**	**8**	**7**	**8**	**7**	**8**	**4**		**57**

Mis-classification table

Results: Classification quality metric ROC-AUC



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	2
ENSEMBLE	M-TSC	0.995	**0.961**	0.002	0.958	5
STSF	U-TSC	0.990	**0.931**	0.004	0.925	1
RIC	U-TSC	0.985	**0.927**	0.006	0.920	3
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-TS	0.979	**0.901**	0.002	0.899	0
STC	U-TS	0.973	**0.897**	0.004	0.892	7
CTSF	U-TS	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	0
QDA	S-CML	-	**0.573**	0.000	0.573	0

MUSE

Conclusions

- Three diverging methods (S-CML, U-TSC, and M-TSC) were proposed and investigated to conduct a **retrospective personalized supervised sport activity classification** of a single person
- A total of **twenty different machine learning models** from *sklearn* and *sktime* were adopted, which makes the study quite extent application of CML models into a specific type of sport data.
- The observed results are mainly good (up to 96,6%), and even better than the average results (92%) in recent HAR studies (Demrozi et al., 2020)
- The selection of the classification method for sport activities depends on the available data, and therefore one model cannot be preferred over others in all the cases
- **MUSE** (96,6%) can be gently proposed **as the best choice** for sport activity classification when using **multivariate time series data**
- However, in certain circumstances, for example **TSF** as a **fast model** could be more preferable than **MUSE**, and then solving this problem in **interlaced univariate data** 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

Discussion...

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?