

Supplementary material for the paper titled “Reliable early classification of time series based on discriminating the classes over time”

Effect of using cross-validation vs. using the pre-set
train/test validation for the ECDIRE method

Usue Mori, Alexander Mendiburu, Eamonn Keogh and
Jose A. Lozano

As stated in Section 5.5 of the paper, the databases from the UCR are provided with pre-defined training and testing sets. As such, in order to enable reproducibility and simplify the comparison between methods, it is common in the literature to use these training and testing sets.

However, in order to show the stability and robustness of ECDIRE, we have carried out some additional experiments using a 10 times repeated 5-fold cross validation procedure, following the recommendations of Rodríguez et al. [1]. For these additional experiments, due to time limitations, we have randomly chosen 10 databases from the UCR archive, obtaining the following results:

Dataset	Acc.(TT)	Acc.(CV)	Earl.(TT)	Earl.(CV)
CBF	0.89	0.99	26.64	49.63
Cricket_X	0.58	0.66	46.83	49.64
Cricket_Z	0.60	0.66	46.95	49.50
Gun_Point	0.92	0.96	32.47	74.08
Haptics	0.45	0.47	71.21	53.50
ItalyPowerDemand	0.93	0.97	67.90	98.01
Lighting7	0.53	0.55	27.11	22.89
MedicalImages	0.73	0.76	32.96	36.26
Trace	0.82	0.88	41.10	45.14
uWaveGestureLibrary_Z	0.71	0.78	74.84	87.29

Table 1: Results for accuracy (Acc.) and earliness (Earl.) for the Train/Test (TT) validation framework and the 10x5-cross validation procedure (CV).

In terms of dominance, the results obtained by the cross validation procedure dominate the train/test methodology in two cases (Haptics and Lighting7), but are equivalent for the remaining cases. Furthermore, in terms of absolute values, the results are very similar for most datasets in both objectives. The accuracy tends to be higher in the case of the CV, resulting in lower earliness values, especially in the databases with high levels of shifting or warping (for example CBF or Gun_Point), or very small predefined training sets, such as CBF. However, no overall large differences are observed and the method is able to provide accurate results early in time.

References

- [1] Juan D. Rodríguez, Aritz Pérez, and Jose A. Lozano. A general framework for the statistical analysis of the sources of variance for classification error estimators. *Pattern Recognition*, 46(3):855–864, 2013.