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

p-values for the statistical tests performed in the paper

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

In this document, we provide all the p-values of the statistical tests included in the main manuscript. First we provide the p-values obtained from the Friedman's tests for each of the objectives separately:

Objective	p-value
Accuracy	8.92e-07
Earliness	4.116e-08

Table 1: p-values obtained from the Friedman test.

Additionally, we also provide the p-values obtained from the pairwise comparisons carried out using the Holm post-hoc test.

	ECDIRE	RelClass	ECTS	EDSC	1NN
ECDIRE	-	0.24	0.13	2.23e-04	0.27
RelClass	-	-	7.12e-03	1.14e-06	0.95
ECTS	-	-	-	0.03	8.74e-03
EDSC	-	-	-	-	1.62e-06
1NN	-	-	-	-	-

Table 2: Raw p-values for the pairwise comparisons of the accuracy.

	ECDIRE	RelClass	ECTS	EDSC	1NN
ECDIRE	-	0.72	0.52	1.78e-03	0.72
RelClass	-	-	0.05	1.14e-05	0.95
ECTS	-	-	-	0.15	0.05
EDSC	-	-	-	-	1.46e-05
1NN	-	-	-	-	-

Table 3: Corrected p-values for the pairwise comparisons of the accuracy, using the Holm post-hoc test.

	ECDIRE	RelClass	ECTS	EDSD
ECDIRE	-	2.38e-05	3.45e-05	0.80
RelClass	-	-	0.93	7.49e-06
ECTS	-	-	-	1.11e-05
EDSD	-	-	-	-

Table 4: Raw p-values for the pairwise comparisons of the earliness.

	ECDIRE	RelClass	ECTS	EDSD
ECDIRE	-	9.52e-05	1.04e-04	1.00
RelClass	-	-	1.00	4.49e-05
ECTS	-	-	-	5.54e-05
EDSD	-	-	-	-

Table 5: Corrected p-values for the pairwise comparisons of the earliness, using the Holm post-hoc test.

Finally, we also provide the p-values obtained for the statistical comparison of the dominance counts using the sign test proposed by Putter [3] and as explained in [1, 4]. This statistical test is a type of sign test with good power properties [2, 1, 4] for cases when there are ties which are not due to rounding errors or the use of discrete surrogate variables [5]. Note that the reported p-values correspond to the one-sided test ($H0 : p+ \leq p-$), which is usually proposed in the literature. However, the two-tailed tests also find significant differences in all cases, for a significance level of 0.05.

	RelClass	ECTS	EDSD
ECDIRE	0.0062	7.996e-06	0.0005

Table 6: p-values for the comparison of the dominance counts.

The p-values are very small and are also found significant if the Holm post-hoc test is applied.

References

[1] Coakley, C., Heise, M., 1996. Versions of the sign test in the presence of ties. Biometrics 52, 1242–1251.

[2] Krauth, J., 1978. An asymptotic UMP sign test in the presence of ties. The Annals of Statistics 6 (2), 461–464.

[3] Putter, J., 1955. The treatment of ties in some nonparametric tests. The Annals of Mathematical Statistics 26 (3), 368–386.

[4] Randles, R., 2001. On neutral responses (zeros) in the sign test and ties in the wilcoxon-mann-whitney test. The American Statistician 55 (2), 96–101.

[5] Wittkowski, K. M., Coakley, C. W., Heise, M. A., 1998. Versions of the sign test in the presence of ties. Biometrics 54 (2), 789–791.