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

Analysis of the effect of modifying the granularity of E and $perc\_acc$ in ECDIRE

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

In this document we analyze the performance and behavior of ECDIRE when we modify two of its parameters: the granularity of E and the value of $perc\_acc$.

1 Effect of modifying the granularity of E

As commented in Section 3.1 of the paper, the first step in the learning process consists in training a set of classifiers within a 10x5-cross validation process that will enable the analysis of the evolution of the accuracy over time. For this, we first select a grid of timestamps E, which will be the only ones that will be considered throughout the process.

In this section, we show the effect of modifying the granularity of E by performing some additional experiments using 6 databases from UCR. The rest of the parameters of ECDIRE are set to the same values used in the paper. The datasets of the UCR contain series of the same length, so we will define the granularities as a percentage of the length of the series in the database, as done in the paper. Note that, since we consider granularities starting from 1% of the length of the series, we have specifically selected databases that contain series of lengths larger than 100.

The results for accuracy and earliness are available in Tables 1 and 2, respectively.

	gran=1%	**gran=2%**	**gran=5%**	**gran=10%**
CBF	0.83	0.88	0.89	0.90
Coffee	0.96	0.96	0.96	0.96
Cricket_X	0.56	0.58	0.59	0.60
FaceFour	0.58	0.58	0.59	0.64
fish	0.81	0.81	0.79	0.82
MedicalImages	0.58	0.58	0.74	0.73

Table 1: Results for accuracy for different values of $perc\_acc$

	gran=1%	gran=2%	gran=5%	gran=10%
CBF	25.38	26.93	28.57	30.11
Coffee	79.32	79.71	82.14	84.29
Cricket_X	42.16	42.84	47.90	55.71
FaceFour	22.21	20.46	22.72	26.92
fish	54.26	54.70	54.89	57.54
MedicalImages	3.22	4.19	21.25	24.89

Table 2: Results for accuracy for different values of $perc_acc$

We also show the results graphically in Figure 1.

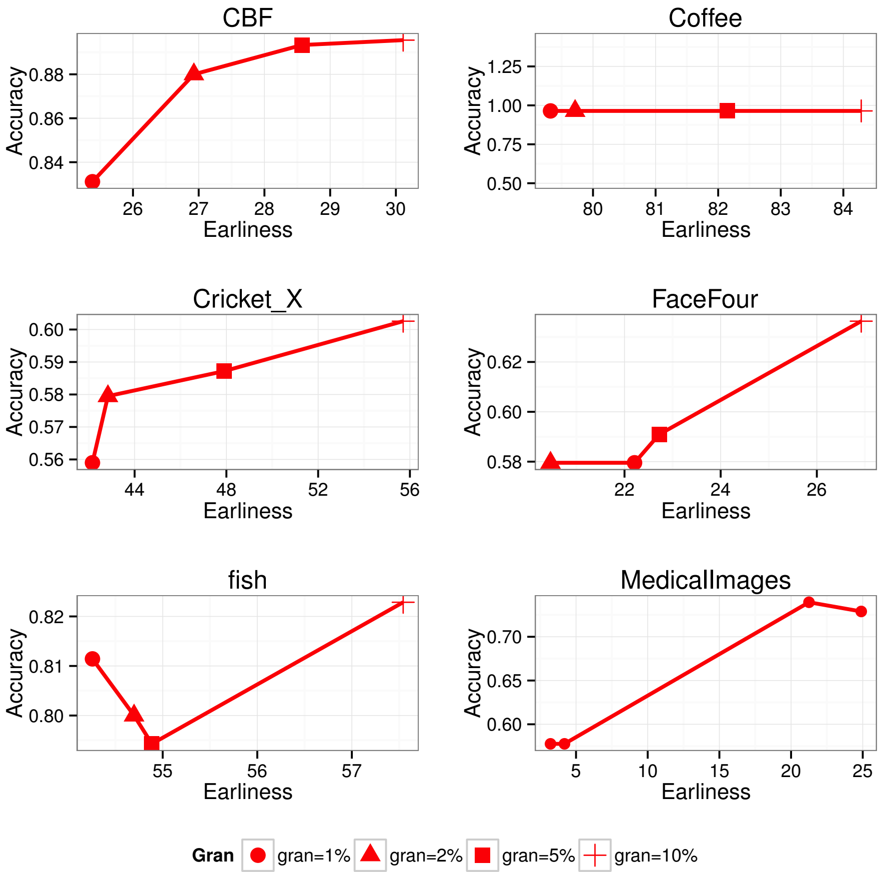
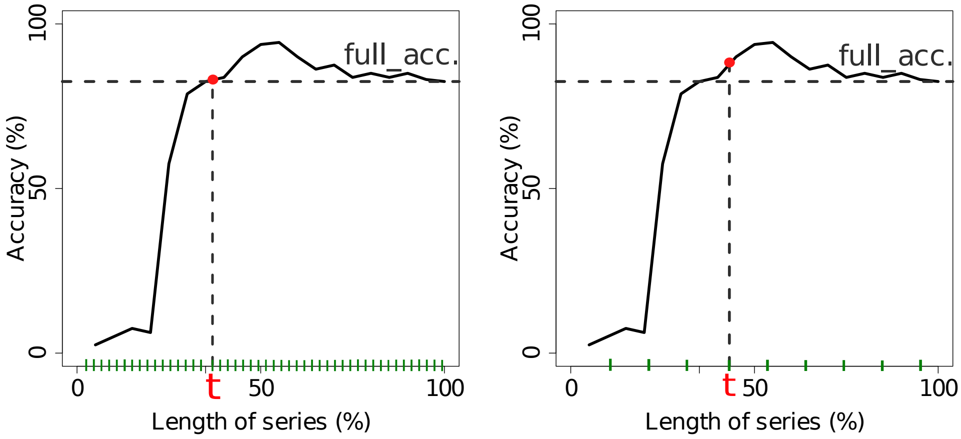


Figure 1: Evolution of earliness and earliness as the granularity of E is modified.

As can be seen, as the granularity is reduced, generally, the earliness tends to decrease. This is expected, to some extent, because we are considering more options when we choose the timestamp in which the accuracy for a class begins to trespass a certain threshold. We can see an example of this in Figure 2. When we consider a finer granularity, the class predictions can be made earlier maintaining the same accuracy threshold.



(a) Granularity 1

(b) Granularity 2.

Figure 2: Selection of the point in the timeline for two different granularities. The granularity considered in each case is shown in the x-axis in green.

Accordingly, as we can see in the figure, the accuracy can also decrease as we increase the granularity. In most cases, this decrease is not very large, because the accuracy requirements are equal for all granularities considered. However, this depends on the database at hand, on the training set we are using, and especially on the evolution of the accuracy over time for each specific database.

In this context, as stated in the paper, choosing an optimal granularity for each database beforehand could be an interesting future research line.

2 Effect of modifying the *perc_acc* parameter

In this section, we analyze the effect of modifying the parameter *perc_acc* for the ECDIRE method. As explained in the article, this parameter represents the level of the accuracy that we consider suitable, and it is defined as a percentage of accuracy that is obtained using the full time series in the training set.

Based on this definition, we expect that, as we lower this parameter from 100%, we will obtain lower accuracy values and, thus, earlier predictions in most cases. We performed some experiments using the 6 databases selected in the previous section obtaining the following results:

	80%	**85%**	**90%**	**95%**	**100%**
CBF	0.77	0.89	0.89	0.89	0.89
Coffee	0.50	0.71	0.89	0.93	0.96
Cricket_X	0.54	0.55	0.56	0.56	0.57
FaceFour	0.52	0.58	0.58	0.61	0.61
fish	0.73	0.80	0.79	0.77	0.81
MedicalImages	0.69	0.69	0.68	0.69	0.74

Table 3: Results for accuracy for different values of *perc_acc*

	80%	**85%**	**90%**	**95%**	**100%**
CBF	23.26	26.69	26.67	26.67	28.55
Coffee	10.36	15.36	57.50	68.57	82.14
Cricket_X	27.41	29.10	32.12	35.42	47.98
FaceFour	16.94	18.79	20	21.21	22.31
fish	25.37	29.86	30.83	35.26	55.17
MedicalImages	6.61	6.68	7.34	9.72	21.20

Table 4: Results for earliness for different values of *perc_acc*

For easier interpretation, we also represent these results graphically in Figure 3. As suspected, in general terms, as the parameter *perc_acc* is lowered, lower accuracy results are obtained, but we classify the series earlier in time.

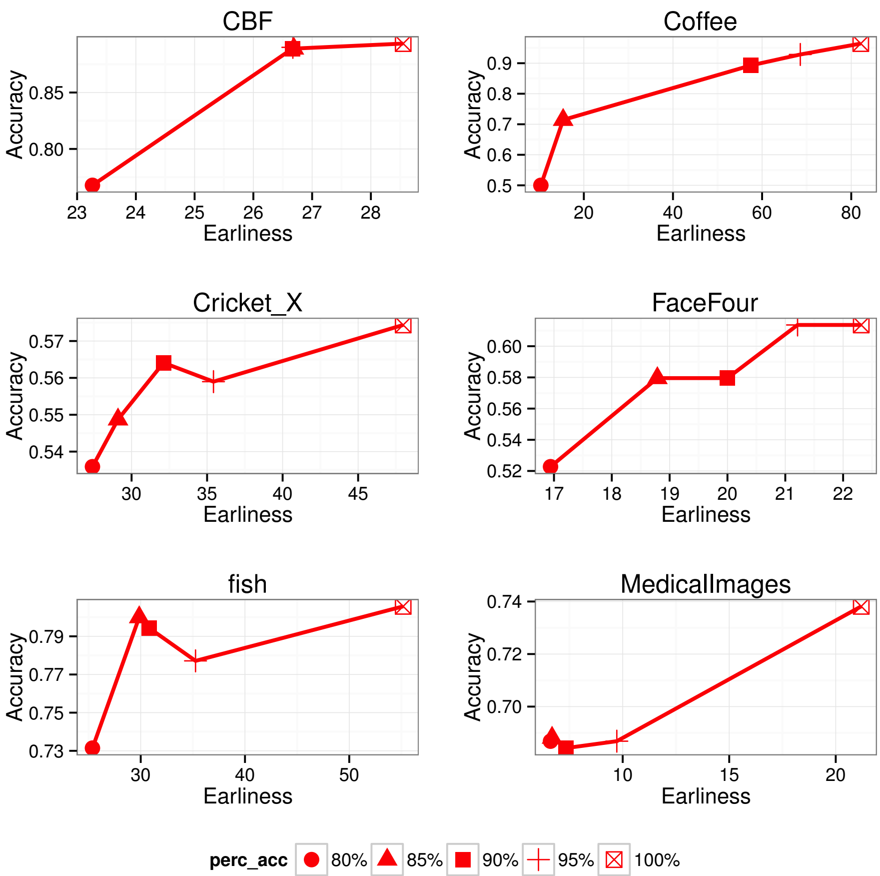


Figure 3: Evolution of accuracy and earliness as the *perc_acc* parameter is modified.

This parameter can be chosen by the user, taking into account the requirements of earliness and accuracy.

In the case of the *coffee* dataset, note that when the *perc_acc* parameter is lowered, the accuracy drops notably. The reason for the difference between the theoretically expected drop in accuracy with the observed one is probably due to the extremely small size of the training dataset (only 28 series) which renders the estimation of the accuracy over time not robust enough. As such, in special cases such as this one, small changes in the *perc_acc* parameter can result in large changes in the timeline, which may not be generalizable to other testing datasets. In these cases we recommend working with a larger training set or using another validation framework different to the 10x5 cross-validation when estimating the accuracy.