

Errata

Learning the parameters of a non compensatory sorting model

Olivier Sobrie

January 10, 2017

1 Binarization of the CEV datasets

In the original document, the binarization of the CEV dataset has been done by thresholding the categories at the median. We compare our results with the ones obtained in Ali Fallah Tehrani, Weiwei Cheng, Krzysztof Dembczynski, and Eyke Hüllermeier. Learning monotone nonlinear models using the Choquet integral. *Machine Learning*, 89(1-2):183–211, 2012. However, in this paper, the binarization is done in another way: the alternatives assigned in the three best categories are grouped together. We provide here the results obtained with the MR-Sort metaheuristic by thresholding in this way.

Size	MR-Sort	NCS	LP UTADIS
20%	0.8652 ± 0.0145	0.8758 ± 0.0142	0.8700 ± 0.0143
50%	0.8612 ± 0.0088	0.8736 ± 0.0081	0.8676 ± 0.0117
80%	0.8585 ± 0.0063	0.8723 ± 0.0048	0.8664 ± 0.0167
100%	0.8576 ± 0.0000	0.8716 ± 0.0000	0.8680 ± 0.0000

Table 1: Classification accuracy of the learning set for difference size of the learning set. These results have been obtained with the binarized CEV dataset.

Size	MR-Sort	NCS	LP UTADIS
20%	0.9057 ± 0.0136	0.9033 ± 0.0152	0.8700 ± 0.0143
50%	0.9018 ± 0.0081	0.9057 ± 0.0103	0.8676 ± 0.0117
80%	0.9001 ± 0.0041	0.9051 ± 0.0060	0.8664 ± 0.0167
100%	0.8996 ± 0.0000	0.9070 ± 0.0000	0.9412 ± 0.0000

Table 2: Area under the curve of the learning set for difference size of the learning set. These results have been obtained with the binarized CEV dataset.

Size	MR-Sort	NCS	LP UTADIS
20%	0.8512 ± 0.0135	0.8650 ± 0.0062	0.8700 ± 0.0142
50%	0.8517 ± 0.0095	0.8664 ± 0.0090	0.8676 ± 0.0117
80%	0.8494 ± 0.0166	0.8683 ± 0.0183	0.8664 ± 0.0167

Table 3: Classification accuracy of the test set for difference size of the learning set. These results have been obtained with the binarized CEV dataset.

Size	MR-Sort	NCS	LP UTADIS
20%	0.8968 ± 0.0116	0.8956 ± 0.0130	0.9235 ± 0.0183
50%	0.8960 ± 0.0073	0.9001 ± 0.0119	0.9339 ± 0.0138
80%	0.8941 ± 0.0135	0.9020 ± 0.0155	0.9399 ± 0.0111

Table 4: Area under the curve of the test set for difference size of the learning set. These results have been obtained with the binarized CEV dataset.