

Table 1: Test errors for the chosen methods on all datasets (↓). **Data Type** indicates the number and types of attributes the dataset has (c is continuous and d is discrete). **Size** denotes the number of data points. In square bracket (if present) next to every dataset name, we provide its acronym referred to in our main text. In the bracket next to each dataset name is either the OpenML ID of the dataset or a reference to the dataset’s associated publication. Error represents one standard deviation.

Dataset	Data Type	Size	Zero-shot	Few-shot	Summary	Summary Boosting
caesarian [cae] (42901)	1c4d	80	0.425± 0.04	0.388± 0.02	0.350± 0.04	0.300± 0.04
iris (61)	4c0d	150	0.680± 0.02	0.460± 0.01	0.275± 0.07	0.193± 0.03
tae (48)	1c4d	151	0.556± 0.07	0.494± 0.01	0.474± 0.02	0.454± 0.03
glass (41)	9c0d	214	0.486± 0.01	0.473± 0.01	0.466± 0.02	0.370± 0.02
breast-cancer [bc] (13)	7c5d	277	0.754± 0.02	0.516± 0.02	0.337± 0.02	0.288± 0.02
visualizing-environmental [ve] (678)	3c0d	111	0.522± 0.01	0.366± 0.01	0.304± 0.02	0.268± 0.03
analcatdata-chlamydia [ac] (535)	2c2d	100	0.200± 0.00	0.200± 0.00	0.170± 0.01	0.170± 0.01
wine (43571)	13c0d	178	0.820± 0.03	0.674± 0.02	0.475± 0.01	0.320± 0.01
blood-transfusion-center [btc] (1464)	4c0d	748	0.544± 0.01	0.430± 0.00	0.258± 0.04	0.240± 0.04
somerville-happiness-survey [shs] [Koczkodaj, 2018]	0c7d	143	0.416± 0.03	0.385± 0.03	0.422± 0.02	0.350± 0.02
vehicle (54)	18c0d	846	0.765± 0.00	0.560± 0.01	0.510± 0.02	0.410± 0.04
statlog-heart [stath] [Dua and Graff, 2017]	6c7d	270	0.551± 0.01	0.528± 0.01	0.444± 0.05	0.430± 0.01
vertebra-column [vc] (1524)	6c0d	310	0.714± 0.03	0.435± 0.06	0.327± 0.01	0.262± 0.01
ecoli (1011)	7c0d	336	0.581± 0.02	0.562± 0.01	0.480± 0.01	0.270± 0.03
haberman-survival [hs] (43)	3c0d	306	0.308± 0.02	0.262± 0.01	0.277± 0.01	0.250± 0.01
diabetes [dia] (37)	8c0d	768	0.446± 0.04	0.400± 0.00	0.360± 0.01	0.344± 0.01
visualizing-hamster [hams] (708)	5c0d	73	0.464± 0.03	0.481± 0.05	0.360± 0.02	0.207± 0.00
wholesale-customers [wc] (1511)	6c1d	440	0.364± 0.01	0.347± 0.01	0.349± 0.02	0.330± 0.00

Table 2: Test errors for chosen methods on all datasets (↓). (acronym and notation in Table 1)

Dataset	Data Type	Size	Summary Boosting	LIFT	KNN	TabPFN	Xgboost
cae (42901)	1c4d	80	0.300± 0.04	0.312± 0.02	0.300± 0.00	0.425± 0.07	0.412± 0.05
iris (61)	4c0d	150	0.193± 0.03	0.100± 0.01	0.106± 0.02	0.027± 0.00	0.054± 0.04
tae (48)	1c4d	151	0.454± 0.03	0.480± 0.04	0.532± 0.01	0.450± 0.13	0.464± 0.01
glass (41)	9c0d	214	0.370± 0.02	0.218± 0.02	0.294± 0.03	0.158± 0.05	0.254± 0.05
bc (13)	7c5d	277	0.288± 0.02	0.318± 0.01	0.277± 0.02	0.264± 0.01	0.270± 0.01
ve (678)	3c0d	111	0.268± 0.03	0.430± 0.04	0.308± 0.01	0.370± 0.04	0.279± 0.02
ac (535)	2c2d	100	0.170± 0.01	0.180± 0.06	0.170± 0.01	0.090± 0.01	0.110± 0.04
wine (43571)	13c0d	178	0.320± 0.01	0.065± 0.01	0.214± 0.05	0.040± 0.01	0.040± 0.01
btc (1464)	4c0d	748	0.240± 0.04	0.270± 0.01	0.238± 0.00	0.209± 0.01	0.219± 0.01
shs [Koczkodaj, 2018]	0c7d	143	0.350± 0.02	0.419± 0.02	0.326± 0.03	0.392± 0.00	0.406± 0.00
vehicle (54)	18c0d	846	0.410± 0.04	0.111± 0.16	0.636± 0.01	0.178± 0.01	0.260± 0.00
stath [Dua and Graff, 2017]	6c7d	270	0.430± 0.01	0.122± 0.17	0.244± 0.03	0.148± 0.03	0.215± 0.00
vc (1524)	6c0d	310	0.262± 0.01	0.192± 0.03	0.318± 0.02	0.135± 0.00	0.187± 0.04
ecoli (1011)	7c0d	336	0.270± 0.03	0.126± 0.03	0.211± 0.03	0.036± 0.02	0.066± 0.01
hs (43)	3c0d	306	0.250± 0.01	0.314± 0.03	0.278± 0.00	0.262± 0.02	0.281± 0.02
dia (37)	8c0d	768	0.344± 0.01	0.324± 0.04	0.353± 0.02	0.238± 0.03	0.234± 0.00
hams (708)	5c0d	73	0.207± 0.00	0.334± 0.08	0.528± 0.02	0.328± 0.01	0.411± 0.01
wc (1511)	6c1d	440	0.330± 0.00	0.125± 0.04	0.043± 0.00	0.088± 0.00	0.098± 0.02