

Data set	$d$	NCA		NCA CS		NCA CS BACK		NCA O1
		1-NN	NCA cls	1-NN	NCA cls	1-NN	NCA cls	1-NN
iris	2	94.44	94.11	95.22	94.78	95.22	95.11	95.56
	3	95.44	95.00	95.33	95.00	94.89	94.22	93.33
	$D = 4$	96.56	96.44	95.56	95.67	95.22	93.11	95.56
wine	2	96.67	96.57	97.04	96.94	97.22	97.59	96.30
	3	96.39	96.39	96.48	95.93	97.22	97.50	98.15
	4	96.20	96.48	98.06	98.06	98.24	98.89	–
	5	96.57	96.57	96.20	95.00	97.78	97.87	–
	$D = 13$	96.20	96.20	95.09	92.59	94.63	98.43	98.15
balance	2	92.74	92.29	87.90	89.04	89.92	91.04	91.49
	3	94.81	94.87	89.60	90.08	65.19	87.85	96.28
	$D = 4$	95.16	95.32	90.82	92.07	67.15	86.91	96.28
ionosphere	2	86.23	86.18	82.83	85.33	85.90	84.20	85.33
	3	87.69	87.64	85.42	83.44	88.11	84.91	79.00
	4	87.64	87.64	86.08	83.58	88.73	83.68	–
	5	89.39	89.39	88.87	81.46	86.65	82.08	–
	$D = 33$	87.97	87.97	86.56	66.09	87.26	84.76	89.39
glass	2	53.62	54.15	55.77	57.23	58.46	59.38	61.54
	3	61.85	60.92	63.15	56.31	61.54	63.46	64.62
	4	62.69	62.15	62.00	54.54	65.23	65.23	–
	5	64.54	63.46	64.92	57.08	67.54	67.23	–
	$D = 9$	66.46	65.54	68.92	57.77	68.15	64.08	63.08

Table 1: Comparison in terms of accuracy of variants of different NCA objective functions on some small data sets. The results are averaged over 30 repetitions, except for the results for NCA O1 which are collected after one experimentation. Abbreviations: NCA represents the simple NCA function, NCA CS denotes the version with compact support kernels, NCA CS BACK denotes the version that uses a background distribution and NCA O1 is the approximate version of simple NCA. Furthermore, 1-NN and NCA cls denote how the classification was done: either using a nearest neighbour rule or by using the function that was also used for learning.