

Appendix A1 for Manuscript BIOINF-2005-1602

Classifying three microarray data using K-nearest neighbor

- Remarks: 1. The average and STD of prediction accuracy are calculated based on 100 iterations of 3-fold cross validation.
 2. For each data set, the range of tested K is determined by the sample size and the number of folds of the cross validation.
 3. The best performance for each data set is highlighted.

LGMD			Leukemia			CNS cancer					
K	Average	STD	K	Average	STD	K	Average	STD	K	Average	STD
1	37.05%	13.04%	1	85.25%	9.16%	1	85.46%	4.76%	31	74.29%	8.45%
2	36.54%	13.09%	2	85.08%	9.04%	2	85.56%	4.90%	32	74.05%	8.42%
3	38.03%	13.37%	3	86.47%	9.13%	3	86.12%	4.81%	33	73.82%	8.39%
4	38.65%	13.37%	4	87.49%	9.07%	4	86.59%	4.65%	34	73.61%	8.36%
5	38.91%	13.17%	5	87.89%	8.87%	5	86.16%	4.66%	35	73.42%	8.32%
6	39.19%	12.99%	6	41.40%	8.84%	6	85.76%	4.71%	36	73.23%	8.27%
7	39.36%	12.90%	7	88.07%	8.91%	7	85.27%	4.77%	37	73.05%	8.23%
8	39.67%	12.93%	8	88.15%	8.89%	8	84.83%	4.84%	38	72.88%	8.19%
9	39.93%	12.83%	9	87.86%	8.88%	9	84.46%	4.88%	39	72.72%	8.14%
10	40.32%	12.75%	10	87.56%	8.90%	10	84.13%	4.89%	40	72.57%	8.09%
11	40.77%	12.69%	11	87.21%	9.10%	11	83.79%	4.96%	41	72.43%	8.04%
12	41.09%	12.71%	12	86.59%	9.41%	12	83.39%	5.14%	42	72.29%	8.00%
13	41.22%	12.77%	13	85.96%	9.80%	13	82.91%	5.38%	43	72.16%	7.95%
14	41.34%	12.78%	14	85.21%	10.14%	14	82.37%	5.67%	44	72.04%	7.90%
15	41.40%	12.77%	15	84.52%	10.44%	15	81.74%	6.05%	45	71.92%	7.85%
16	41.33%	12.81%	16	83.83%	10.72%	16	81.06%	6.49%	46	71.80%	7.80%
17	41.23%	12.79%	17	83.14%	11.11%	17	80.36%	6.92%	47	71.69%	7.76%
18	40.92%	12.69%	18	82.22%	11.91%	18	79.69%	7.28%	48	71.59%	7.71%
19	40.51%	12.60%	19	80.51%	13.80%	19	79.05%	7.61%	49	71.49%	7.66%
20	40.05%	12.54%	20	78.96%	15.14%	20	78.45%	7.86%	50	71.39%	7.61%
21	39.54%	12.50%	21	77.30%	16.53%	21	77.91%	8.05%	51	71.30%	7.57%
			22	75.80%	17.57%	22	77.40%	8.20%	52	71.21%	7.52%
						23	76.94%	8.31%	53	71.12%	7.48%
						24	76.51%	8.39%	54	71.04%	7.43%
						25	76.11%	8.44%	55	70.96%	7.39%
						26	75.75%	8.48%	56	70.88%	7.34%
						27	75.41%	8.49%	57	70.81%	7.30%
						28	75.10%	8.50%	58	70.74%	7.25%
						29	74.81%	8.49%	59	70.67%	7.21%
						30	74.54%	8.47%	60	70.60%	7.17%

Classification performance on three microarray data sets using one-vs-all support vector machine

Remarks: 1) The average and STD of prediction accuracy are resulted from 100 iterations of 3-fold cross validation.

2) For each data set, the model(s) with best performance is highlighted.

LGMD Penalty	All genes													
	Linear		Poly 2		Poly 3		Gaus 0.01		Gaus 0.1		Gaus 0.5		Gaus 1.0	
	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD
0.001	50.06%	13.14%	15.56%	4.14%	15.53%	4.00%	29.46%	2.70%	29.15%	2.83%	37.50%	13.27%	49.38%	12.77%
0.01	50.25%	12.42%	15.56%	4.14%	15.55%	4.09%	29.46%	2.70%	29.33%	3.54%	38.21%	12.96%	48.63%	13.02%
0.1	50.25%	13.44%	15.56%	4.14%	15.55%	4.09%	29.46%	2.70%	30.15%	5.29%	38.03%	13.48%	48.72%	12.87%
1.0	50.52%	13.08%	15.56%	4.14%	15.57%	4.12%	29.39%	2.62%	30.68%	6.70%	38.13%	13.10%	48.43%	13.66%
10.0	50.08%	13.04%	15.56%	4.14%	15.57%	4.12%	28.85%	2.55%	30.34%	6.50%	42.66%	13.91%	50.94%	12.58%

LGMD Penalty	186 JDG genes													
	Linear		Poly 2		Poly 3		Gaus 0.01		Gaus 0.1		Gaus 0.5		Gaus 1.0	
	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD
0.001	96.77%	5.32%	78.69%	10.97%	82.12%	14.19%	29.39%	2.78%	29.26%	3.23%	90.04%	8.77%	99.43%	2.21%
0.01	100.00%	0.00%	78.69%	10.97%	82.12%	14.19%	29.33%	2.79%	31.87%	7.66%	89.84%	8.38%	99.21%	2.62%
0.1	100.00%	0.00%	78.69%	10.97%	82.92%	13.90%	29.33%	2.79%	37.45%	11.84%	90.07%	8.68%	99.44%	2.23%
1.0	100.00%	0.00%	77.85%	11.50%	82.92%	13.90%	29.35%	2.73%	43.01%	14.73%	89.28%	9.00%	99.94%	0.68%
10.0	100.00%	0.00%	77.85%	11.50%	82.27%	13.27%	29.20%	2.77%	39.02%	13.10%	98.86%	3.51%	100.00%	0.00%

Leukemia Penalty	All genes													
	Linear		Poly 2		Poly 3		Gaus 0.01		Gaus 0.1		Gaus 0.5		Gaus 1.0	
	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD
0.001	93.66%	6.08%	76.49%	11.34%	84.47%	12.83%	44.19%	1.79%	44.19%	1.79%	83.03%	10.57%	88.05%	10.15%
0.01	94.41%	6.12%	75.78%	12.54%	84.47%	12.83%	44.19%	1.79%	44.19%	1.79%	81.59%	10.60%	87.77%	9.52%
0.1	95.10%	6.52%	74.94%	12.55%	85.79%	13.10%	44.19%	1.79%	44.19%	1.79%	83.00%	10.40%	88.66%	8.62%
1.0	95.34%	5.97%	74.94%	12.55%	85.79%	13.10%	44.19%	1.79%	44.19%	1.79%	83.09%	9.59%	89.95%	7.87%
10.0	95.00%	5.97%	76.73%	11.58%	85.18%	12.21%	44.19%	1.79%	44.19%	1.79%	85.89%	9.79%	92.93%	8.06%

Leukemia Penalty	7 JDG genes													
	Linear		Poly 2		Poly 3		Gaus 0.01		Gaus 0.1		Gaus 0.5		Gaus 1.0	
	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD
0.001	95.86%	7.24%	93.56%	7.52%	88.79%	10.58%	44.19%	1.79%	65.84%	9.54%	96.57%	4.58%	96.76%	4.40%
0.01	95.98%	6.78%	94.22%	7.11%	91.23%	8.07%	44.19%	1.79%	65.23%	9.74%	96.41%	4.76%	96.51%	5.48%
0.1	96.57%	4.39%	94.97%	6.72%	93.00%	8.15%	44.19%	1.79%	65.63%	9.80%	95.97%	5.86%	96.79%	4.26%
1.0	96.25%	4.97%	94.94%	6.86%	93.00%	8.15%	44.19%	1.79%	67.40%	10.60%	96.31%	4.79%	96.16%	4.57%
10.0	98.37%	3.76%	94.94%	6.86%	92.48%	8.92%	44.19%	1.79%	72.13%	10.15%	96.61%	5.00%	97.39%	4.23%

CNS cancer Penalty	All genes													
	Linear		Poly 2		Poly 3		Gaus 0.01		Gaus 0.1		Gaus 0.5		Gaus 1.0	
	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD
0.001	89.04%	3.43%	71.38%	3.93%	66.71%	1.69%	66.71%	1.79%	66.72%	1.85%	66.71%	1.72%	66.71%	1.72%
0.01	88.97%	3.59%	71.59%	4.14%	66.71%	1.76%	66.71%	1.78%	66.71%	1.77%	66.71%	1.72%	66.72%	1.82%
0.1	88.82%	3.52%	71.56%	3.87%	66.72%	1.84%	66.72%	1.87%	66.72%	1.91%	66.71%	1.78%	66.72%	1.96%
1.0	89.13%	3.49%	71.41%	4.38%	66.71%	1.77%	66.71%	1.78%	66.72%	1.95%	66.72%	1.88%	74.54%	4.60%
10.0	88.96%	3.38%	71.48%	4.32%	66.72%	1.90%	66.71%	1.78%	66.72%	1.83%	68.32%	2.88%	83.65%	3.81%

CNS cancer Penalty	19 JDG genes													
	Linear		Poly 2		Poly 3		Gaus 0.01		Gaus 0.1		Gaus 0.5		Gaus 1.0	
	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD	Average	STD
0.001	66.72%	1.93%	81.99%	4.21%	84.61%	4.47%	66.71%	1.72%	66.72%	1.87%	66.71%	1.75%	66.71%	1.74%
0.01	83.91%	4.11%	88.68%	4.44%	85.18%	5.30%	66.72%	1.80%	66.71%	1.78%	66.71%	1.73%	66.72%	1.90%
0.1	91.89%	3.33%	88.75%	4.85%	85.49%	4.78%	66.72%	1.90%	66.72%	1.83%	66.72%	1.89%	66.71%	1.78%
1.0	93.92%	3.67%	88.72%	5.03%	85.34%	4.98%	66.72%	1.94%	66.72%	1.82%	88.07%	3.86%	90.60%	2.98%
10.0	94.12%	3.46%	89.01%	4.75%	84.93%	5.26%	66.72%	1.86%	66.71%	1.71%	92.80%	4.05%	95.59%	3.25%