

Random_Intensity_90_1a: 90 Cases, 90 Controls, 300 Peaks

Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis

24 peaks selected as putative biomarkers by the 10 methods within BMDK

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
10	4									
22									5	
57	5									
62		2	3		2		2	5		2
66		3			3		3			3
76	3							1		
77						3				
100		5	1	4		2	5	1		5
104									5	
125			2	3						
145				5		3			3	
147		4			4		4	5		4
161									5	
192	2									
194								5		
216	1	1	3		1		1	3		1
222						3				
223									5	
232									5	
237									3	
239				2					1	
241			3					4		
247				1		1			1	
277					5					

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 24 putative biomarkers.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	63.3	64.8	53.4	61.8	65.8	69.8	67.5	69.0	67.5	69.0
Spec	57.8	70.8	74.7	60.0	73.4	78.6	71.8	74.1	71.8	74.1
%Undet	0.0	1.7	2.8	0.6	13.9	26.1	6.7	15.6	6.7	15.6
Quality	121.1	133.9	125.3	121.2	125.3	122.3	132.6	127.5	132.6	127.6

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 24 putative biomarkers with the caveat that %Undetermined cannot exceed 5.0%.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	63.3	64.8	53.4	61.8	None	None	62.1	None	62.1	None
Spec	57.8	70.8	74.7	60.0	None	None	70.1	None	70.1	None
%Undet	0.0	1.7	2.8	0.6	None	None	3.3	None	3.3	None
Quality	121.1	133.9	125.3	121.2	None	None	128.9	None	128.9	None

Fingerprint Analysis

Sensitivity, specificity and quality (sensitivity + specificity) for the best and 200th best decision tree constructed from any of the 300 peak intensities. The evolutionary programming search used a population size of 200 and ran for 400 generations. A decision node became a terminal node when it contained 1% (no samples) or 4% (3 samples) of a given State.

Metric	1%		1%		4%		4%	
	1 st	200 th						
Sensitivity	76.7	77.8	80.0	80.0	77.8	73.3	86.7	85.6
Specificity	80.0	77.8	80.0	78.9	83.3	85.6	75.6	73.3
Quality	156.7	155.6	160.0	158.9	161.1	158.9	162.2	158.9

Sensitivity, specificity and quality (sensitivity + specificity) for the best and 200th best medoid classifier algorithm in each of the two runs using 5-, 6-, and 7-peak intensities from the set of 300. The evolutionary programming search used a population size of 400 and ran for 800 generations with the requirement that there are at most 60 Case-cells and 60 Control-cells.

Metric	5-Features		5-Features		6-Features		6-Features		7-Features		7-Features	
	1 st	200 th										
Sens	100.0	100.0	84.4	77.8	100.0	100.0	85.6	80.0	100.0	100.0	88.9	82.2
Spec	84.4	78.8	100.0	100.0	88.9	81.1	100.0	100.0	88.9	83.3	100.0	100.0
Quality	184.4	178.8	184.4	177.8	188.9	181.1	185.6	180.0	188.9	183.3	188.9	182.2

(Last updated 4/21/07)