

Random_Intensity_30_1a: 30 Cases, 30 Controls, 300 Peaks

Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis

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

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
6									2	
17						3				
24									2	
28		1	2	2	1		1	2		1
36	5	2	4	4			2	3		2
41						1				
66	4					5		4		
83				5					1	
84					5			4		
94		3	3	3	2		3	4		3
100	3	5			4		4			5
139									2	
143						5				
155								4		
167	2									
207	1	4	1	1	3		5	1		4
208									2	
219			5							
226						4				
257									2	
259						2				

Peaks used in each of the best distance-dependent 6-nearest neighbor classifiers

Peak	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
6										
17										
24										
28	X		X	X			X		X	X
36										
41										
66								X		
83							X		X	
84						X				
94										
100						X				X
139										
143										
155		X	X		X					
167										
207					X	X		X		X
208								X		
219										
226		X	X							
257										
259										

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 21 putative biomarkers.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	70.0	70.0	76.7	72.4	69.2	95.5	85.2	79.2	85.2	88.0
Spec	76.7	82.8	82.8	77.3	88.0	72.7	87.5	84.0	87.5	72.0
%Undet	0.0	1.7	1.7	15.0	15.0	26.7	15.0	18.3	15.0	16.7
Quality	146.7	151.1	157.8	134.7	142.2	141.5	157.7	144.8	157.7	143.3

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 21 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	70.0	70.0	76.7	75.0	None	None	75.0	None	75.0	None
Spec	76.7	82.8	82.8	56.7	None	None	62.1	None	62.1	None
%Undet	0.0	1.7	1.7	3.3	None	None	5.0	None	5.0	None
Quality	146.7	151.1	157.8	128.3	None	None	132.1	None	132.1	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% (1 sample) of a given State.

Metric	1%		1%		4%		4%	
	1 st	200 th						
Sensitivity	100.0	96.7	96.7	96.7	96.7	100.0	96.7	95.0
Specificity	100.0	100.0	100.0	100.0	93.3	90.0	100.0	95.0
Quality	200.0	196.7	196.7	196.7	190.0	190.0	196.7	190.0

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 20 Case-cells and 20 Control-cells.

Metric	5-Features		5-Features		6-Features		6-Features		7-Features		7-Features	
	1 st	200 th										
Sens	100.0	100.0	100.0	93.3	100.0	100.0	100.0	96.7	100.0	100.0	100.0	100.0
Spec	100.0	93.3	100.0	100.0	100.0	96.7	100.0	100.0	100.0	96.7	100.0	100.0
Quality	200.0	193.3	200.0	193.3	200.0	196.7	200.0	196.7	200.0	196.7	200.0	200.0

(Last updated 4/21/07)