

Random_Intensity_30_5a: 30 Cases, 30 Controls, 300 Peaks

Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis

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

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
11	2									
17									5	
28			3	2		2			2	
34				5					4	
108									5	
110	3	1			1		1	2	5	1
124		3			3	5	3	4		3
125	4									
138				5					5	
146									5	
169			1	1		1		2	1	
190		5	5		4		5	4		5
205	1	2	2	4	2	4	2	1		2
218	5									
223			3	2		2			3	
298		4			5		4	4		4

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
11										
17										
28		X			X		X		X	
34										
108										
110			X			X		X		X
124	X	X		X	X	X	X		X	
125			X							
138										
146										
169								X		X
190										
205			X							
218								X		X
223										
298						X				

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 16 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	84.0	70.0	78.3	73.1	80.8	87.5	80.8	91.3
Spec	75.9	76.7	78.6	78.6	92.3	91.3	88.9	74.1	88.9	74.1
%Undet	1.7	0.0	11.7	3.3	18.3	18.3	11.7	15.0	11.7	16.7
Quality	144.2	146.7	150.9	145.2	152.2	146.0	158.0	146.6	158.0	148.7

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 16 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	None	70.0	None	None	82.1	None	82.1	None
Spec	75.9	76.7	None	78.6	None	None	63.3	None	63.3	None
%Undet	1.7	0.0	None	3.3	None	None	3.3	None	3.3	None
Quality	144.2	146.7	None	145.2	None	None	142.1	None	142.1	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	100.0	100.0	93.3	93.3	96.7	93.3
Specificity	96.7	96.7	100.0	100.0	96.7	90.0	100.0	100.0
Quality	196.7	193.3	200.0	200.0	190.0	183.3	196.7	193.3

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	100.0	100.0	100.0	100.0	96.7	100.0	100.0
Quality	200.0	193.3	200.0	193.3	200.0	200.0	200.0	196.7	200.0	100.0	200.0	200.0

(Last updated 4/21/07)