

Random_Intensity_60_1a: 60 Cases, 60 Controls, 300 Peaks

Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis

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

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
2									4	
5								5		
30						4				
56	5									
60		5					5			5
68						4				
79						1				
88		3	2	4	5		4	1		3
110									4	
113			3	5	2			3		
163								5		
193	2									
206						2				
216			4	3					2	
220			5	1					1	
228	3	2			3	4	2	5		2
240						2				
241									2	
252		1	1	2	1	4	1	2		1
260	4									
261						4				
285	1	4			4		3	3	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
2										
5										
30										
56										
60						X				
68										
79										
88										
110										
113			X		X	X		X		X
163										
193										
206						X				
216	X		X		X		X		X	
220										
228		X								
240										
241								X		X
252							X		X	
260			X							
261		X		X						
285								X		X

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 22 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	69.6	70.0	69.8	76.7	69.1	71.7	67.3	71.7
Spec	66.7	68.3	74.6	66.7	66.7	76.9	73.2	71.2	71.9	71.2
%Undet	0.0	0.0	4.2	0.0	13.3	31.7	7.5	12.5	6.7	12.5
Quality	136.7	138.3	140.1	136.7	123.1	122.0	134.8	130.4	132.5	130.4

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 22 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	69.6	70.0	None	None	68.3	None	68.6	None
Spec	66.7	68.3	74.6	66.7	None	None	66.1	None	66.1	None
%Undet	0.0	0.0	4.2	0.0	None	None	3.3	None	3.3	None
Quality	136.7	138.3	140.1	136.7	None	None	131.1	None	131.1	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% (2 samples) of a given State.

Metric	1%		1%		4%		4%	
	1 st	200 th						
Sensitivity	88.3	88.3	88.3	88.3	83.3	83.3	88.3	83.3
Specificity	88.3	86.7	86.7	81.7	86.7	83.3	88.3	88.3
Quality	176.7	175.0	175.0	170.0	170.0	166.7	176.7	171.7

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 40 Case-cells and 40 Control-cells.

Metric	5-Features		5-Features		6-Features		6-Features		7-Features		7-Features	
	1 st	200 th										
Sens	100.0	100.0	88.3	83.3	100.0	100.0	93.3	86.7	100.0	100.0	95.0	90.0
Spec	91.7	83.3	100.0	100.0	91.7	85.0	100.0	100.0	95.0	90.0	100.0	100.0
Quality	191.7	183.3	188.3	183.3	191.7	185.0	193.3	186.7	195.0	190.0	195.0	190.0

(Last updated 4/21/07)