Description of Tables PDF Version

Random_Intensity_300_4a: 300 Cases, 300 Controls, 300 Peaks

Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis

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

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	yip
1								1		
9					1					
12			3	5		4				
15									3	
21						3				
39	5							3		
57					3					
70				2					1	
75									2	
89			5	4						
91	4		4			1				
94						2				
102								3		
103						5				
111					2					
141		3	2	3			3			3
146		5					5			5
162									3	
188						5				
211	3									
228		4	1	1			4			4
235									3	
262	2	2					2			2
264					4					
271					5			3		
274	1	1					1	2		1

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
1					X	X		X	X	
9										
12					X					
15										
21										
39										
57										
70										
75										
89										
91	X	X	X	X		X	X	X		X
94										
102										
103										
111										
141			X							X
146										X
162										
188										
211		X	X			X	X	X	X	
228										
235										
262										
264										
271										
274										

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 26 putative biomarkers.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	56.3	59.0	58.7	56.4	59.4	61.6	59.1	63.2	62.4	60.4
Spec	56.3	59.3	59.0	56.3	61.3	62.5	59.3	60.3	56.7	61.4
%Undet	0.0	0.0	0.0	0.3	5.5	19.2	0.3	11.2	0.3	11.8
Quality	112.7	118.3	117.7	112.3	115.1	106.5	118.1	112.0	118.7	109.9

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 26 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	56.3	59.0	58.7	56.4	55.1	None	59.1	None	62.4	None
Spec	56.3	59.3	59.0	56.3	57.3	None	59.3	None	56.7	None
%Undet	0.0	0.0	0.0	0.3	4.5	None	0.3	None	0.3	None
Quality	112.7	118.3	117.7	112.4	107.9	None	118.1	None	118.7	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 400 and ran for 800 generations. A decision node became a terminal node when it contained 1% (3 samples) or 4% (12 samples) of a given State.

Metric	1%		1%		4%		4%	
	1st	200th	1st	200th	1st	200th	1st	200th
Sensitivity	63.0	67.3	77.7	78.3	74.3	70.3	71.3	70.7
Specificity	73.0	67.7	59.0	57.0	62.0	63.3	64.7	64.0
Quality	136.0	135.0	136.7	135.3	136.3	133.7	136.0	134.7

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 1000 and ran for 2000 generations with the requirement that there are at most 200 Case-cells and 200 Control-cells.

Metric	5-Features		5-Features		6-Features		6-Features		7-Features		7-Features	
	1st	200th	1st	200th	1st	200th	1st	200th	1st	200th	1st	200th
Sens	100.0	100.0	67.0	63.3	100.0	100.0	76.7	71.3	100.0	100.0	78.7	75.0
Spec	68.0	63.3	100.0	100.0	76.7	72.3	100.0	100.0	78.3	75.3	100.0	100.0
Quality	168.0	163.3	167.0	163.3	176.7	172.3	176.7	171.3	178.3	175.3	178.7	175.0

(Last updated 4/21/07)