

Aldehyde Deuteration - Model Search - full dataset

Model Search

formula	McFadden R2
class ~ dip_y + NPA_3 + diff 1 2 + loc.B5	0.633
class ~ dip_y + NPA_1 + NPA_3 + loc.B5	0.631
class ~ -2-3- + dip_y + Dist(2, 7) + NPA_1	0.626
class ~ -2-3- + dip_y + NPA_3 + diff 1 2	0.626
class ~ dip_y + Dist(2, 7) + NPA_3 + diff 1 2	0.615
class ~ Total + NPA_3 + diff 1 2 + loc.B5	0.612
class ~ dip_y + Dist(2, 7) + NPA_1 + NPA_3	0.611
class ~ Total + NPA_1 + NPA_3 + loc.B5	0.611
class ~ Total + diff 1 2 + diff 2 3 + loc.B5	0.610
class ~ dip_y + NPA_3 + diff 1 2 + L	0.607

1st Place

Table:

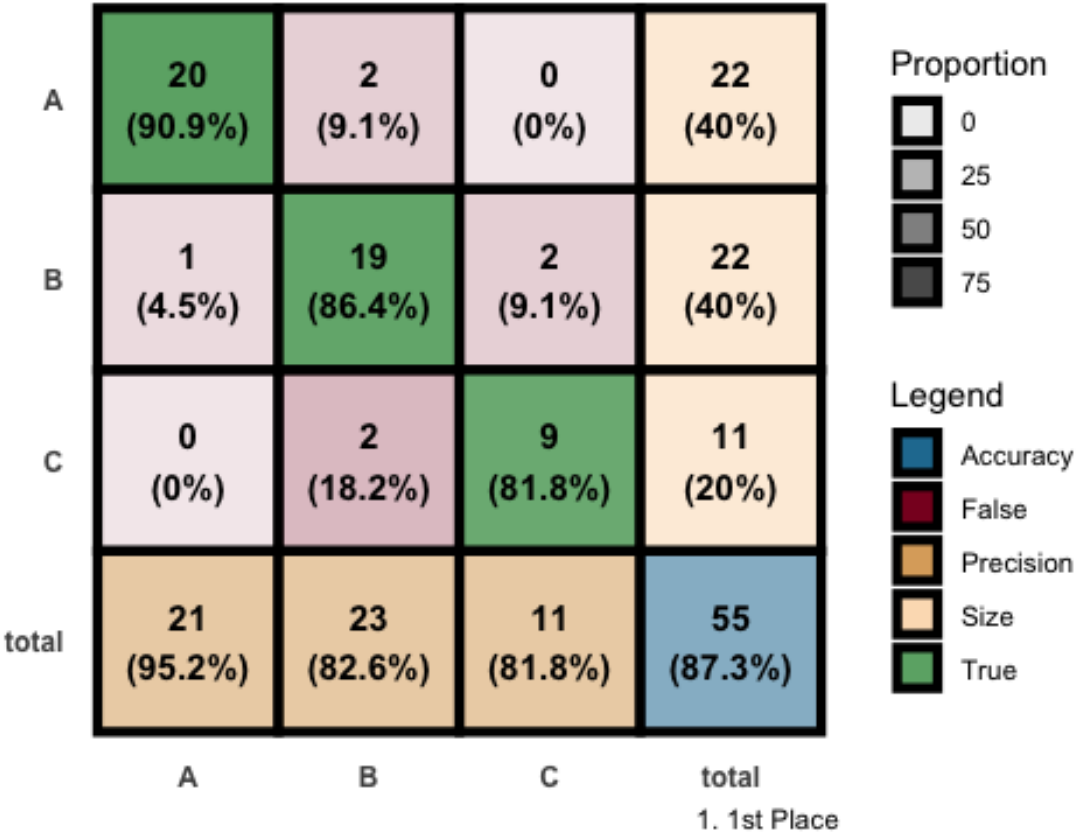
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
87.27%	0.633

Model Coefficients

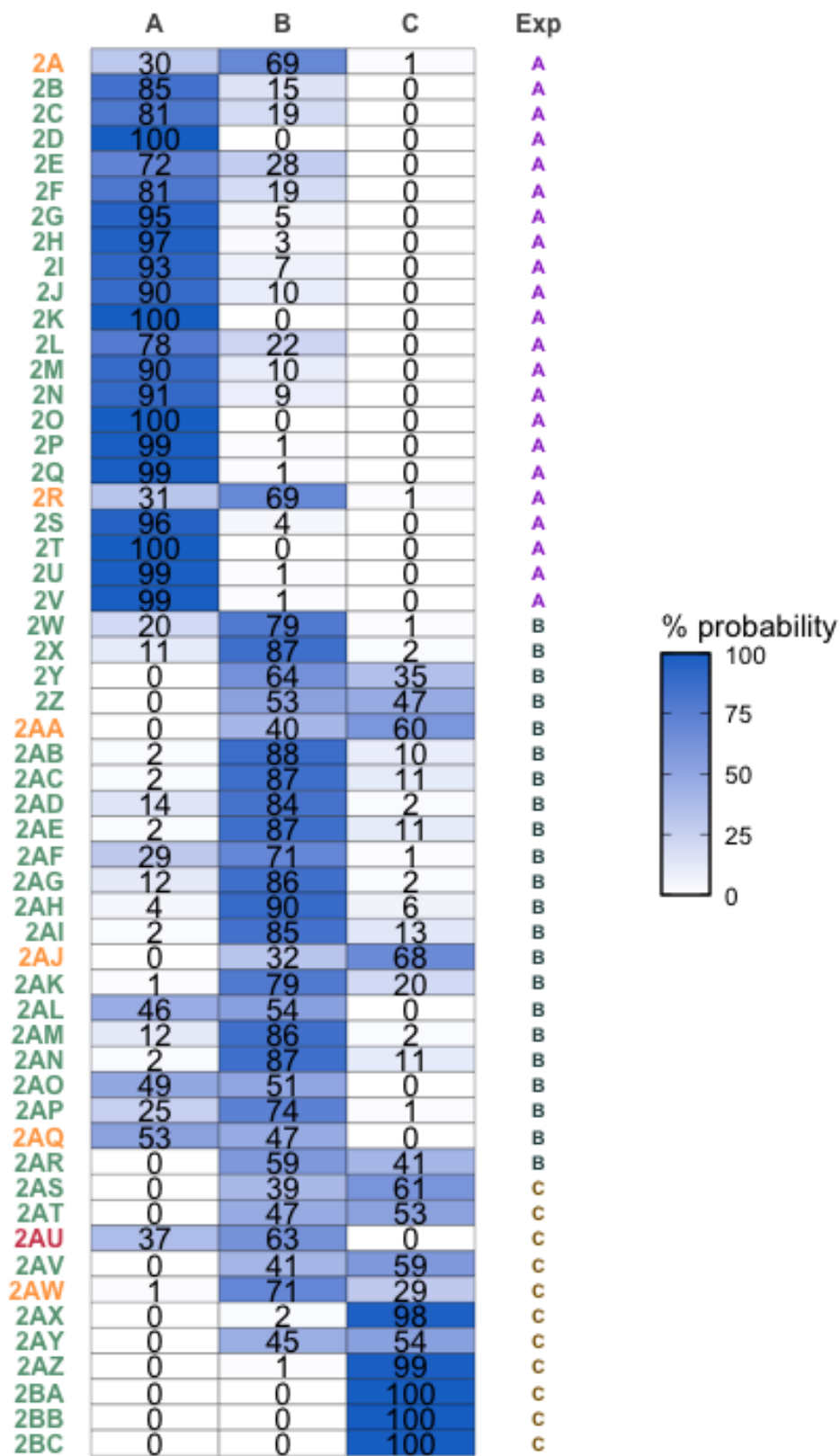
	x
dip_y	-2.1731078
NPA_3	-301.0306563
diff 1 2	-45.1831292
loc.B5	-0.4137891

Training Set
Confusion Matrix



Training Set

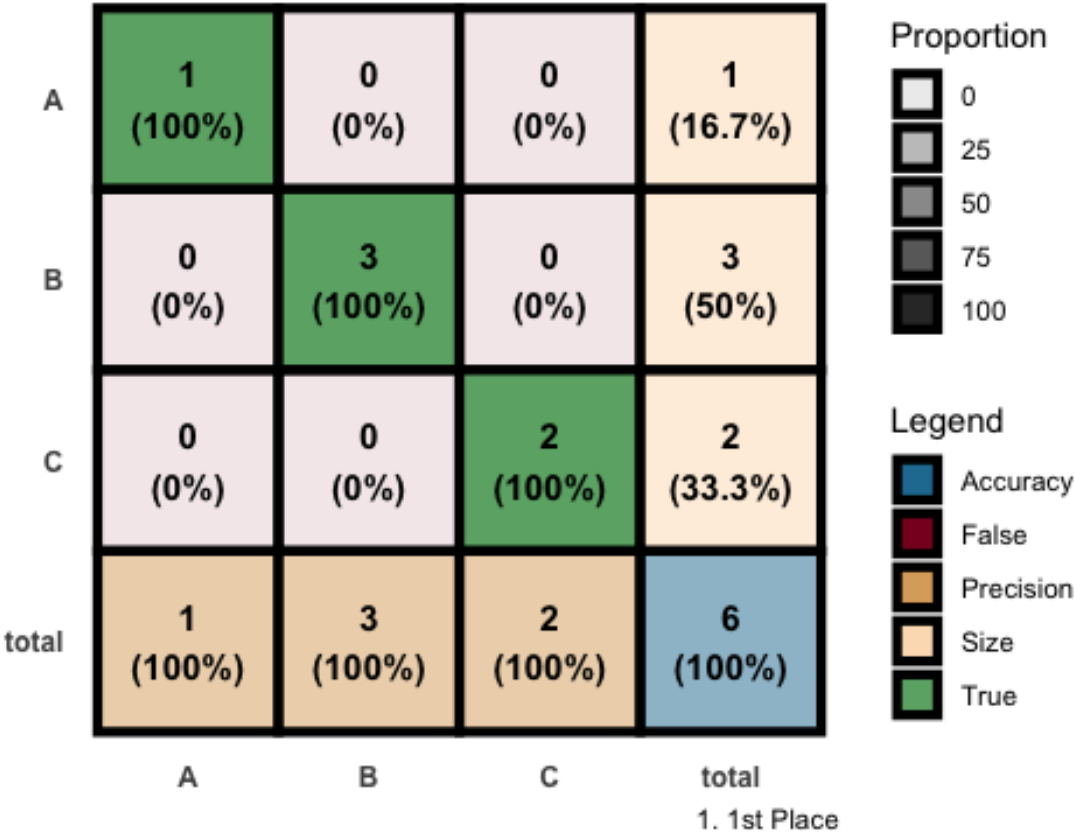
Probability Heatmap



1. 1st Place

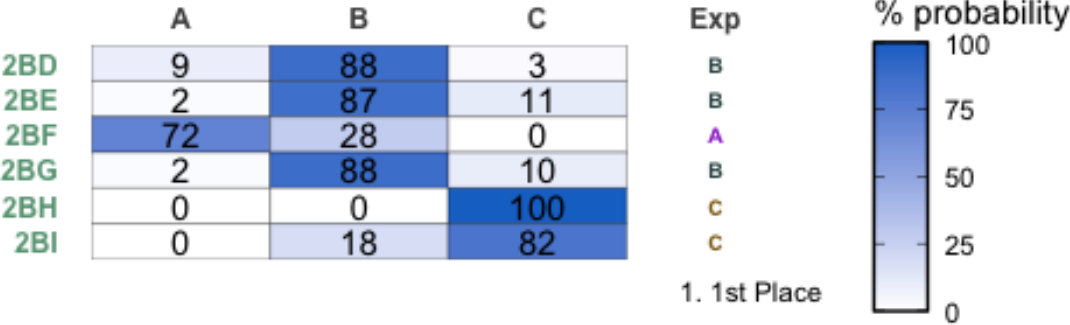
External Validation

Confusion Matrix



External Validation

Probability Heatmap



2nd Place

Table:

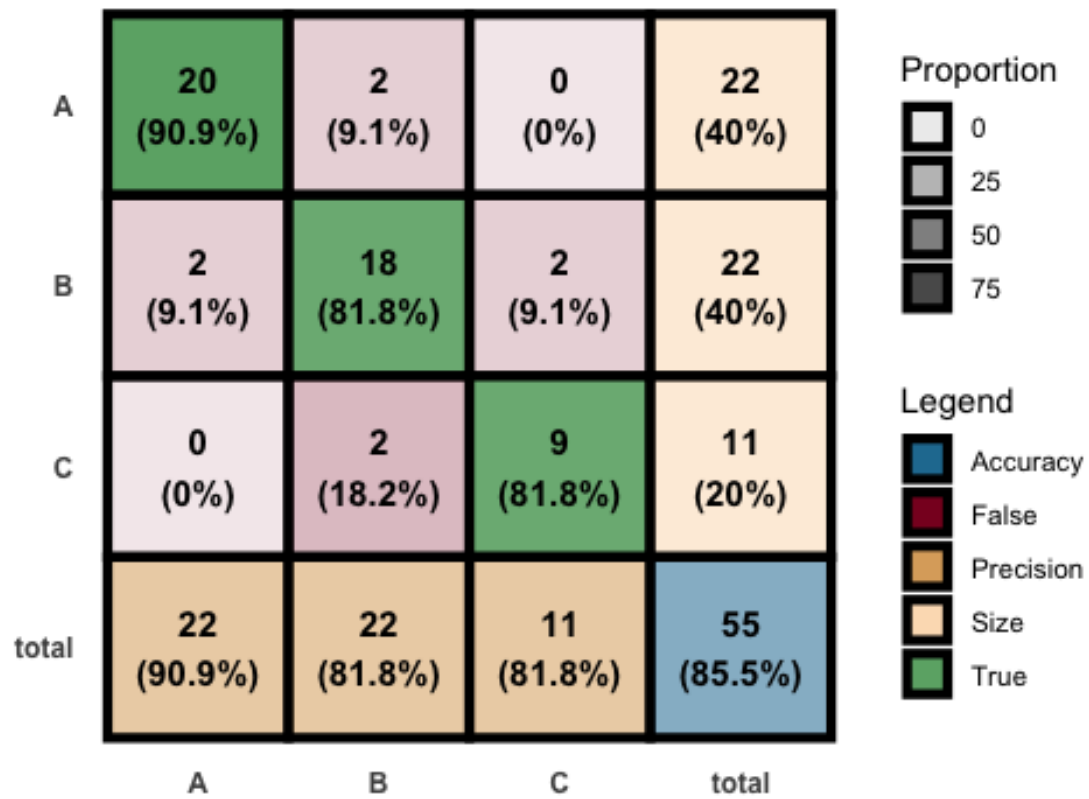
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.45%	0.631

Model Coefficients

	x
dip_y	-2.2317015
NPA_1	44.8723551
NPA_3	-296.7042382
loc.B5	-0.4538779

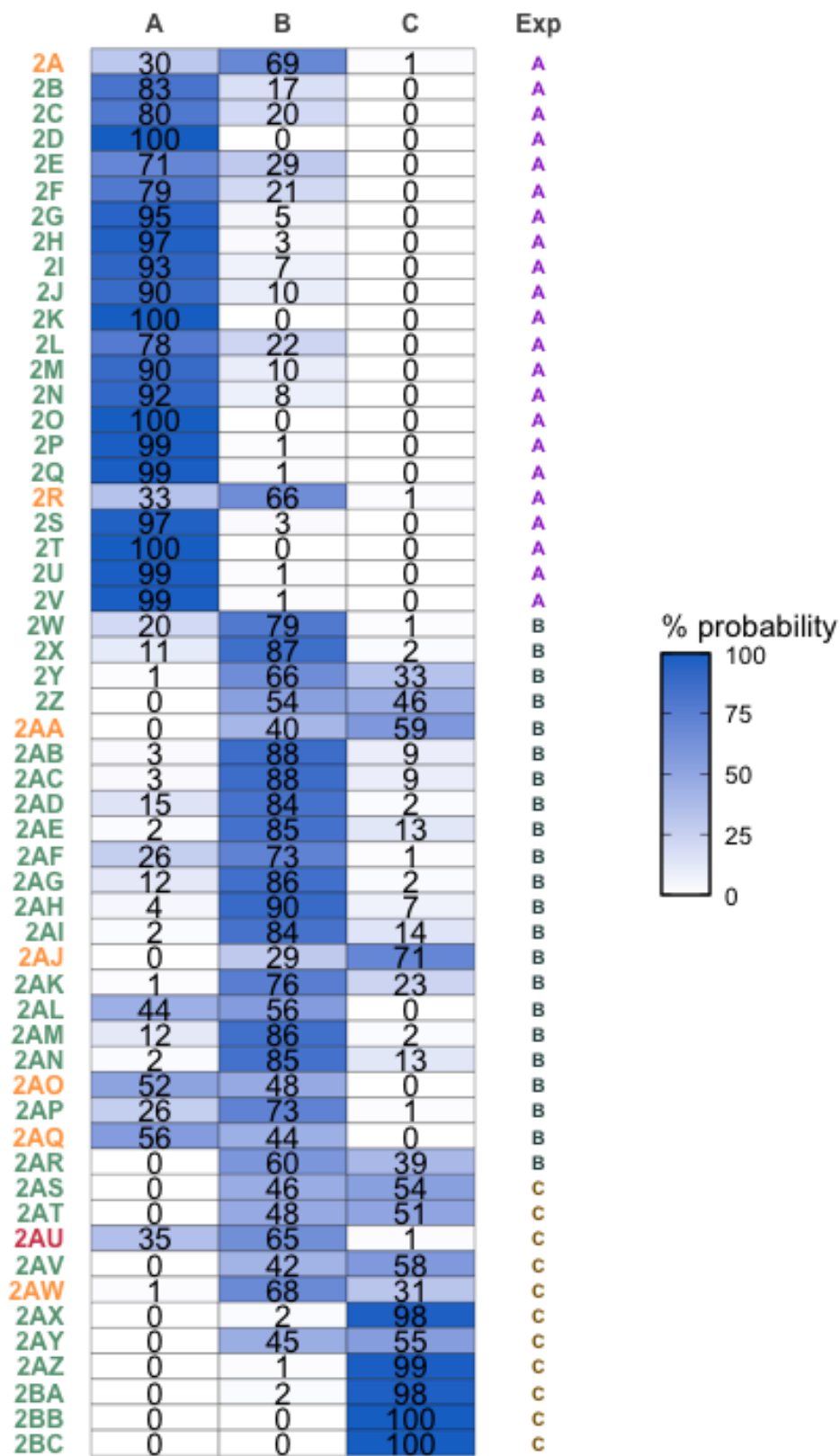
Training Set Confusion Matrix



2. 2nd Place

Training Set

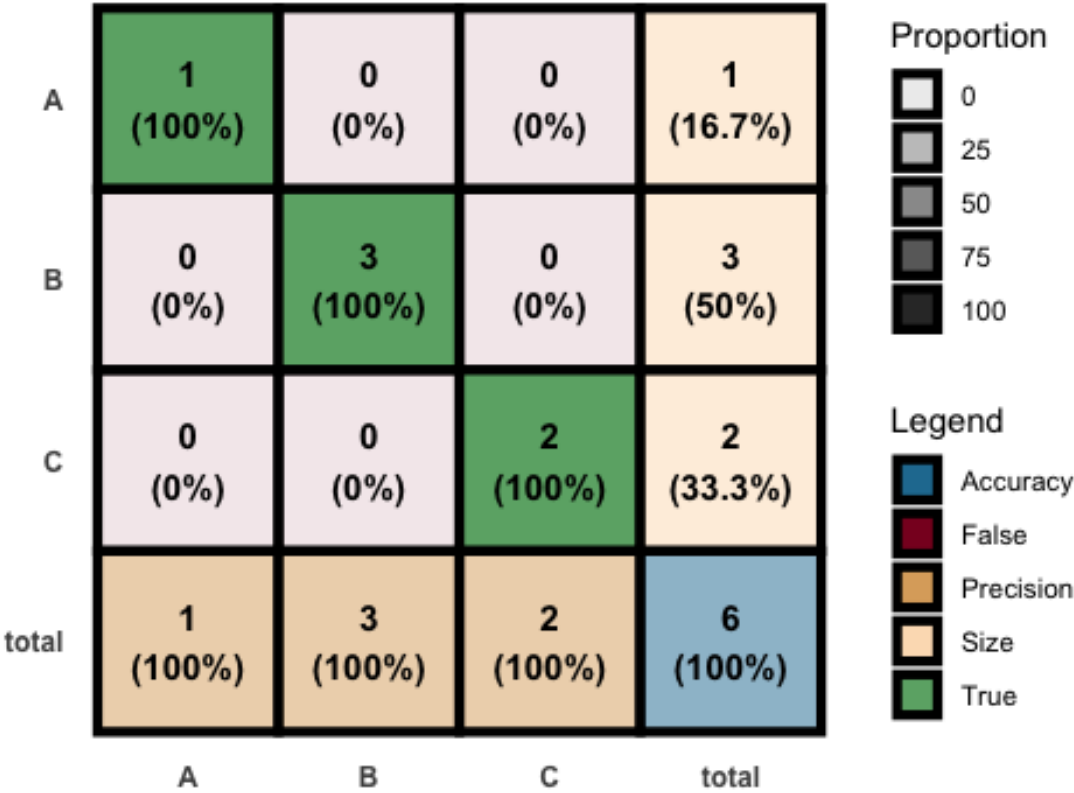
Probability Heatmap



2. 2nd Place

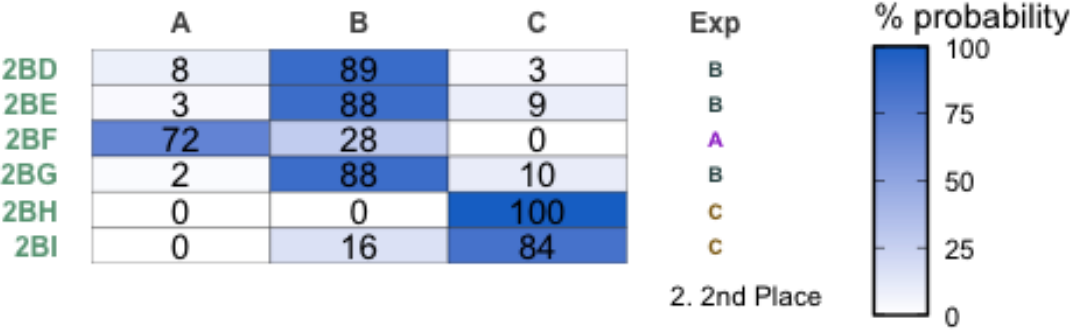
External Validation

Confusion Matrix



External Validation

Probability Heatmap



3rd Place

Table:

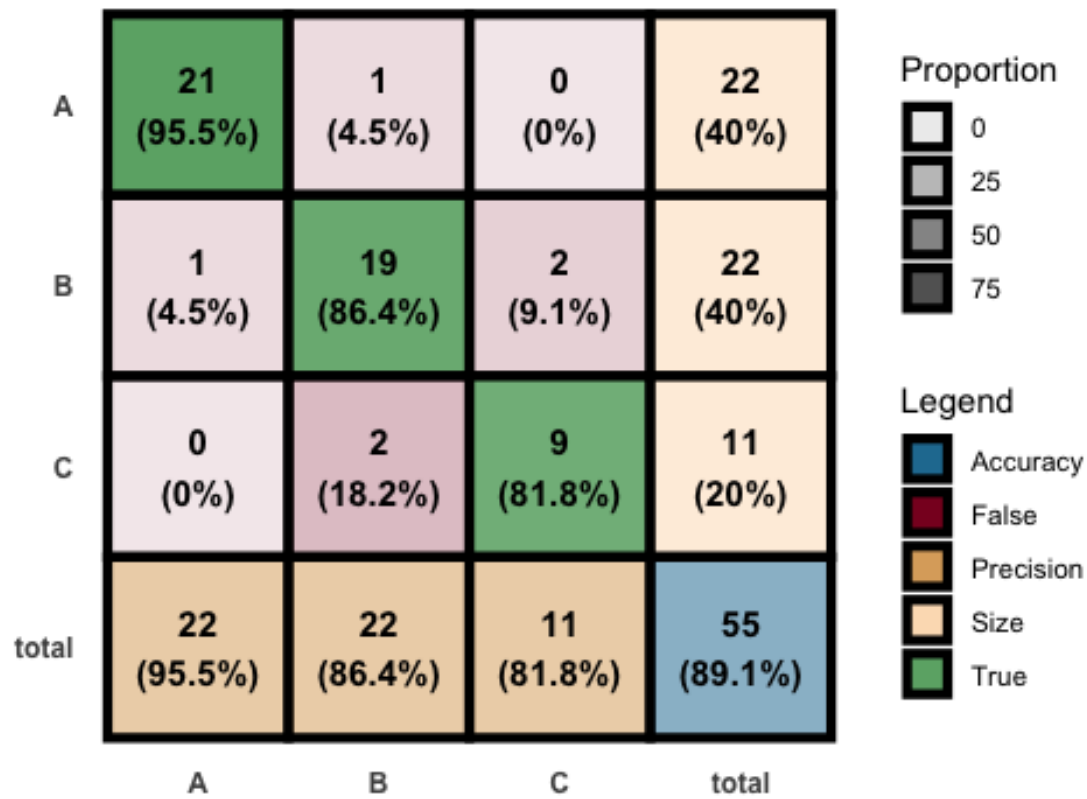
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
89.09%	0.626

Model Coefficients

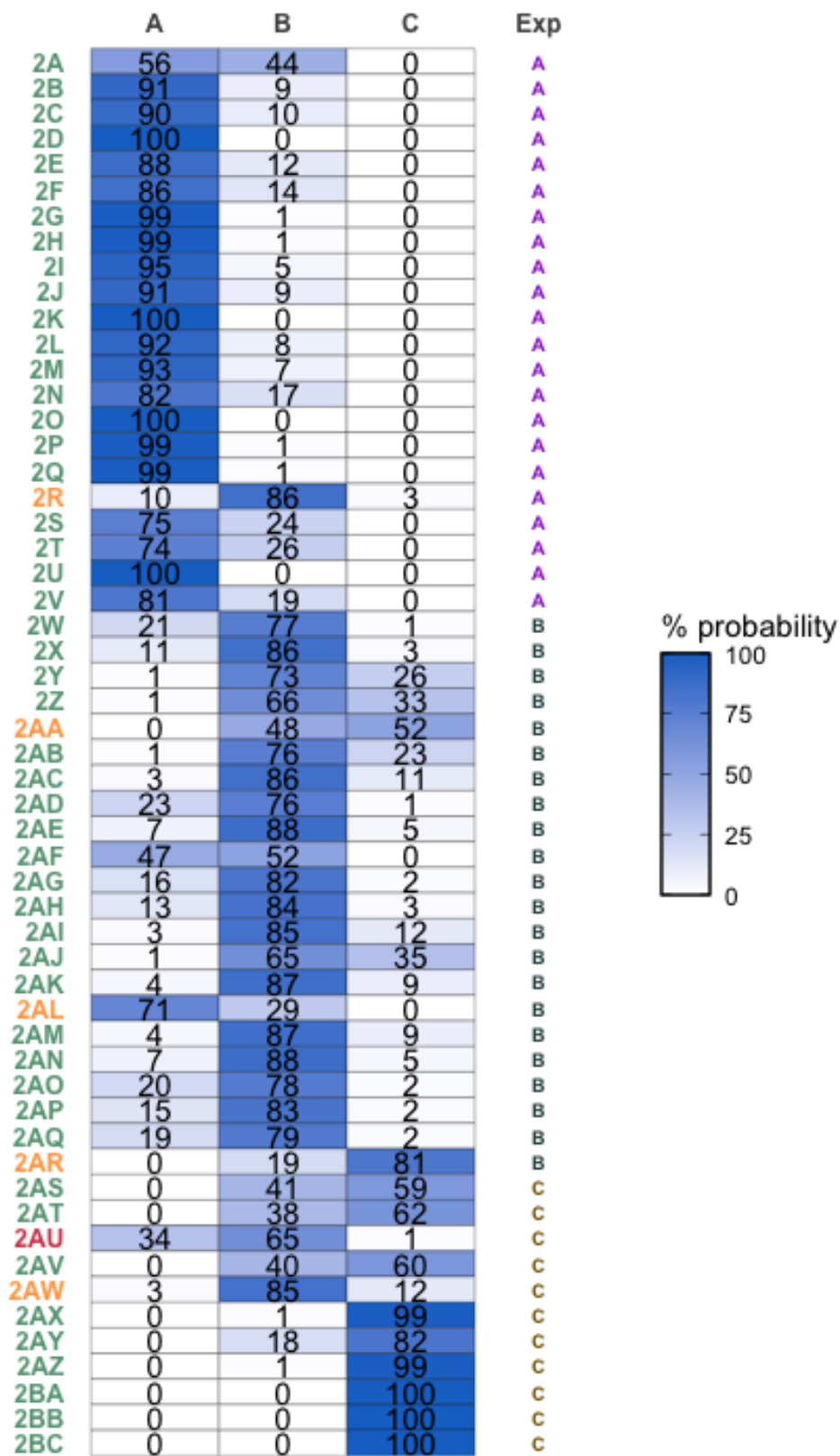
	x
-2-3-	-0.4678155
dip_y	-1.8657764
Dist(2, 7)	349.7731912
NPA_1	56.0996610

Training Set Confusion Matrix



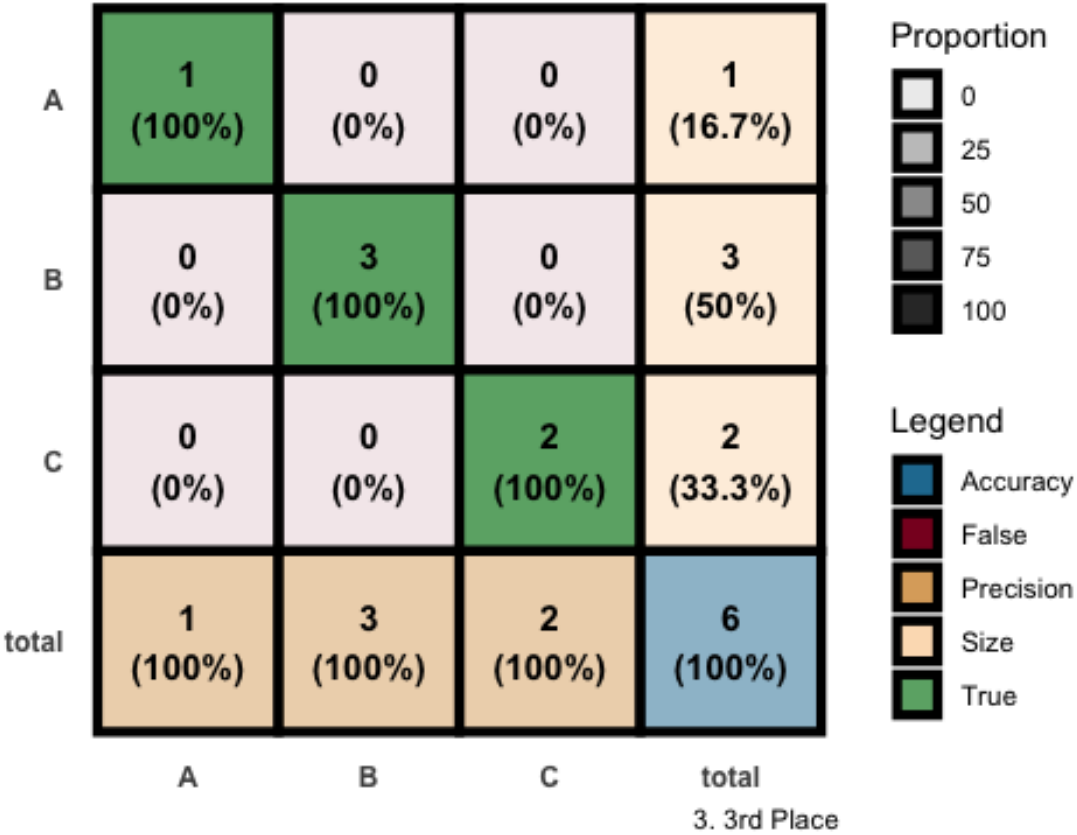
Training Set

Probability Heatmap



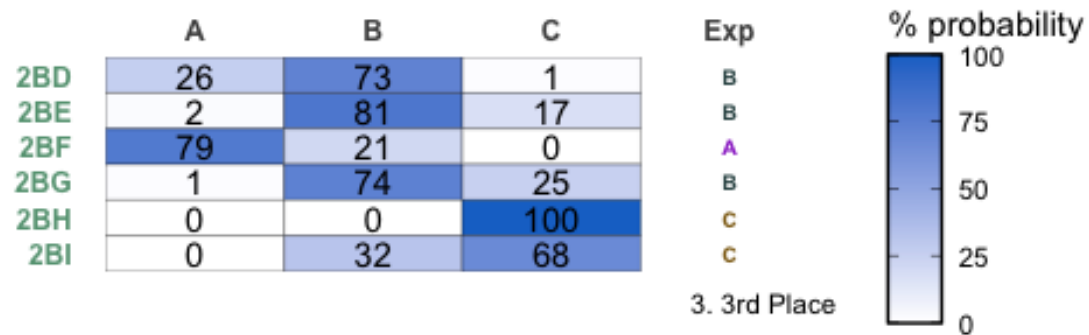
3. 3rd Place

External Validation
Confusion Matrix



External Validation

Probability Heatmap



4th Place

Table:

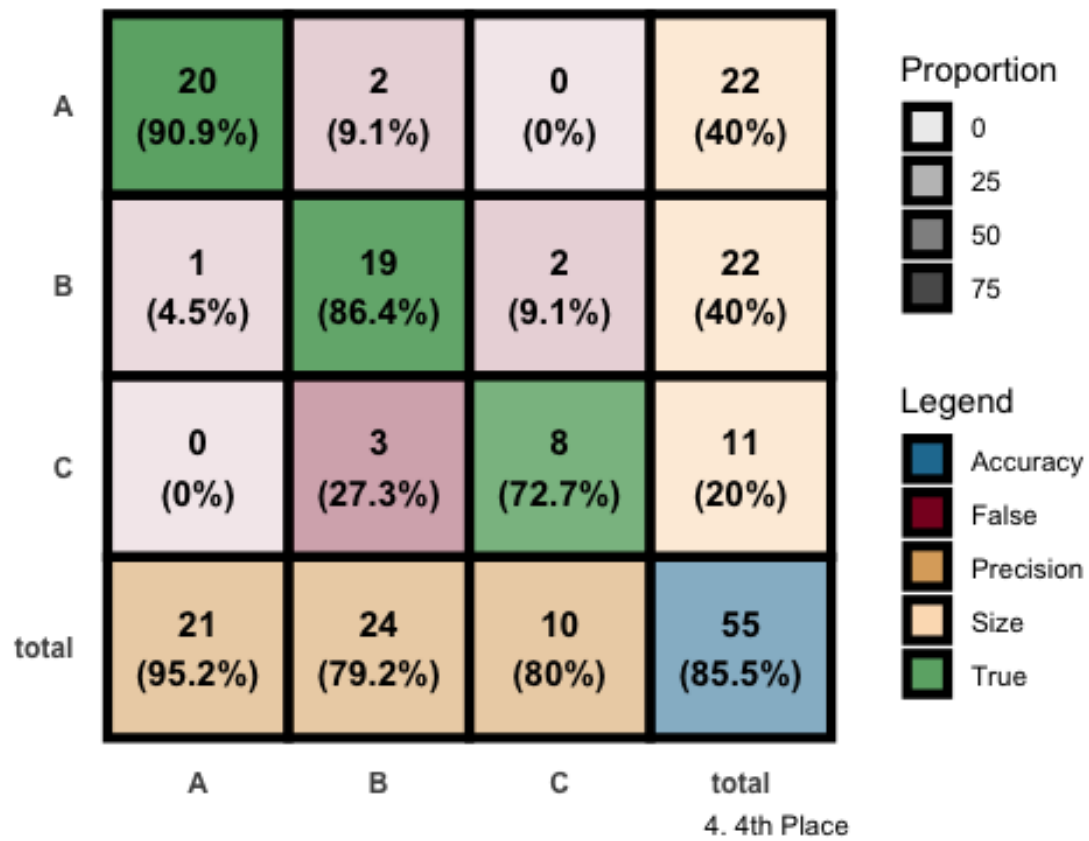
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.45%	0.626

Model Coefficients

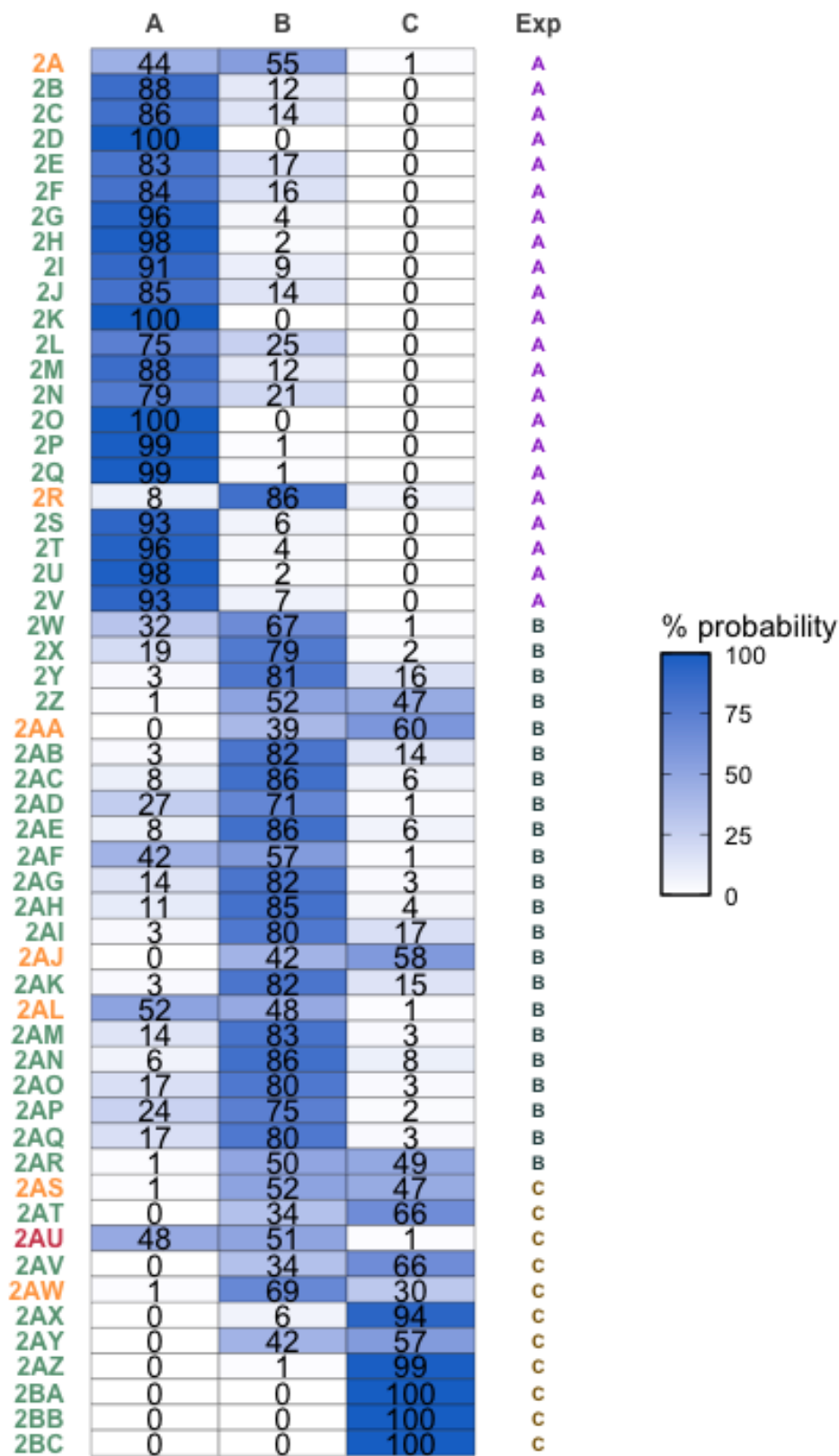
	x
-2-3-	-0.1274129
dip_y	-1.5102414
NPA_3	-229.1140961
diff 1 2	-37.2903965

Training Set Confusion Matrix



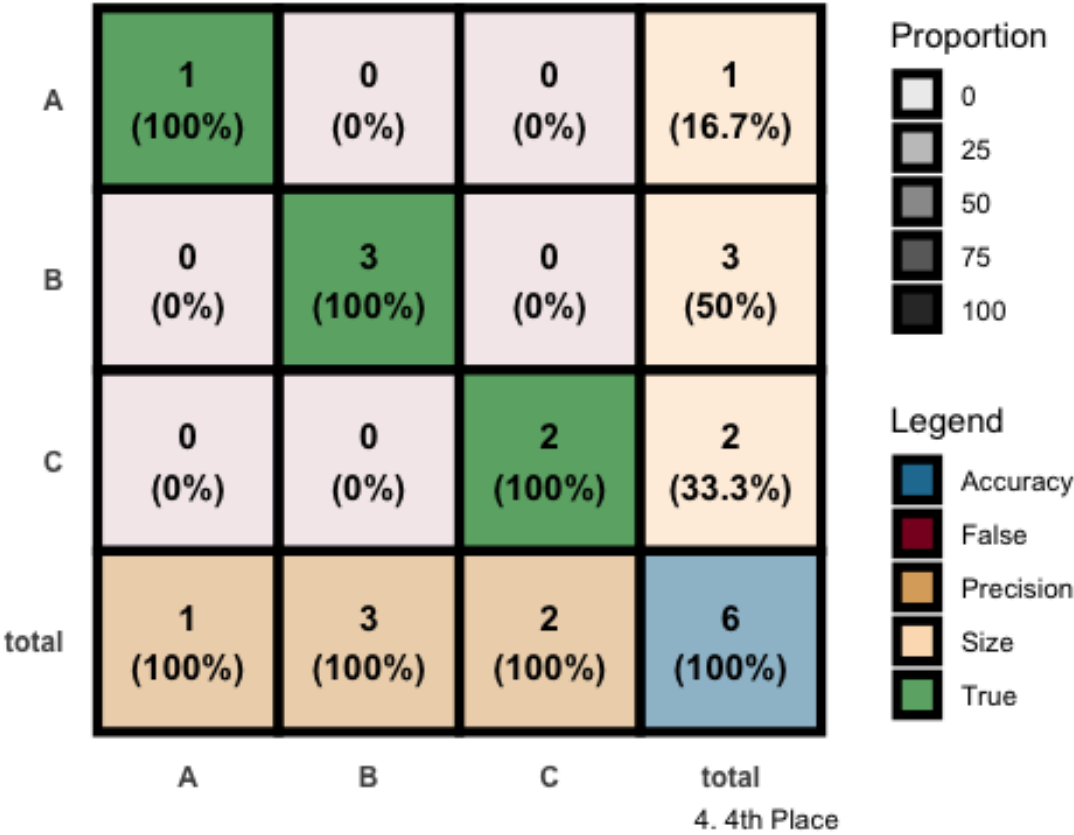
Training Set

Probability Heatmap



External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	20	78	2
2BE	1	67	32
2BF	58	42	0
2BG	2	75	23
2BH	0	0	100
2BI	0	26	74

Exp

B

B

A

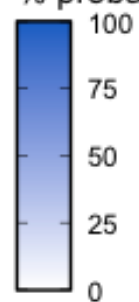
B

C

C

4. 4th Place

% probability



5th Place

Table:

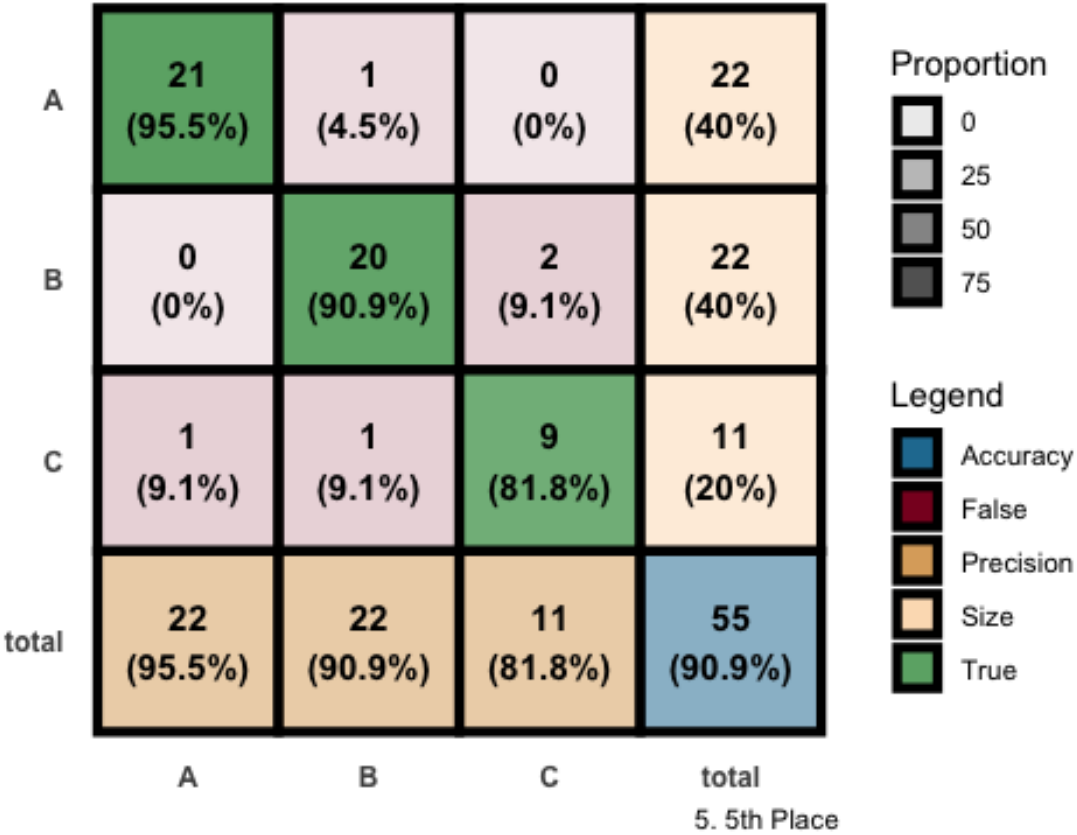
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
90.91%	0.615

Model Coefficients

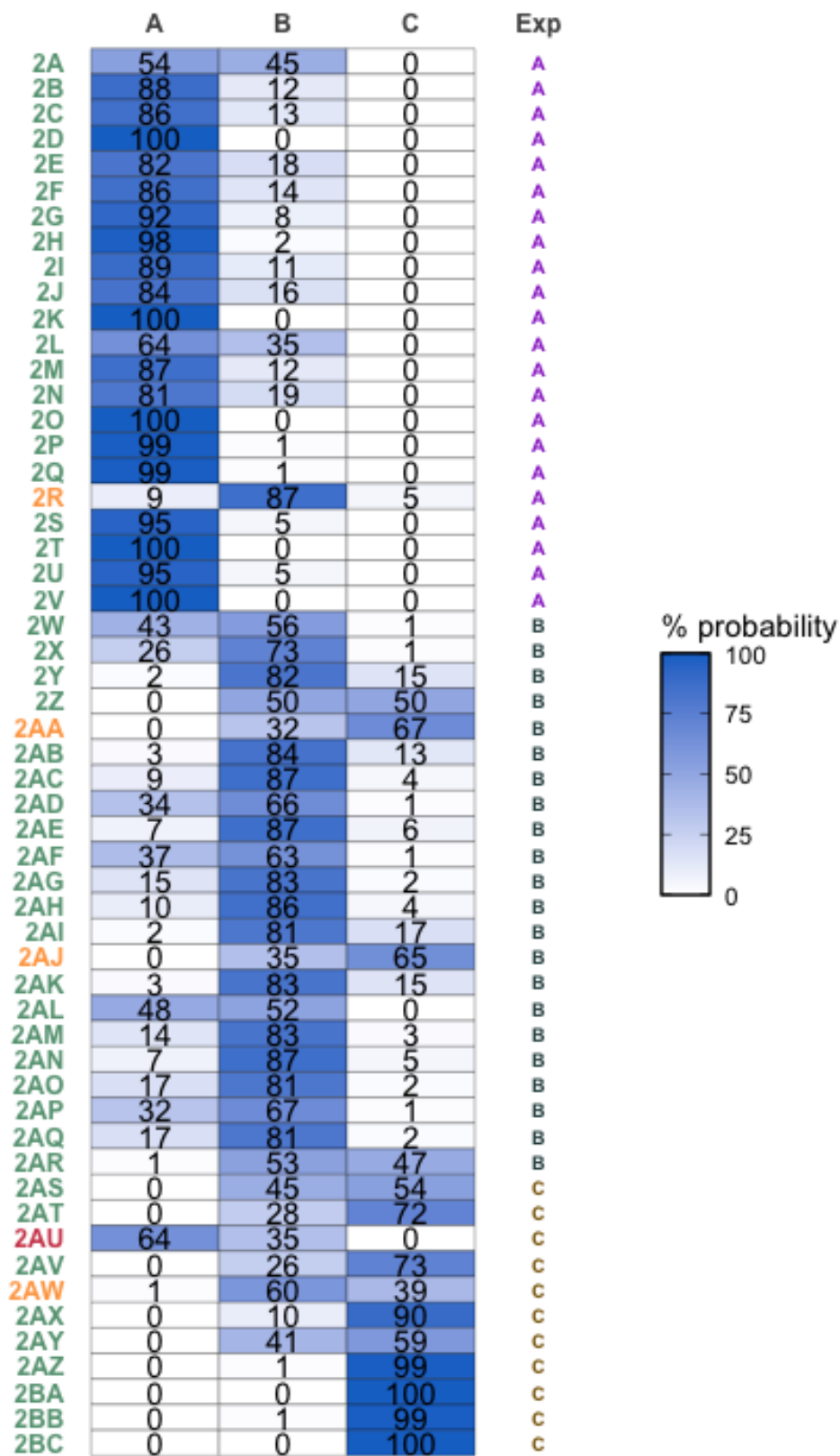
	x
dip_y	-1.312088
Dist(2, 7)	-199.772943
NPA_3	-377.371537
diff 1 2	-40.854377

Training Set
Confusion Matrix



Training Set

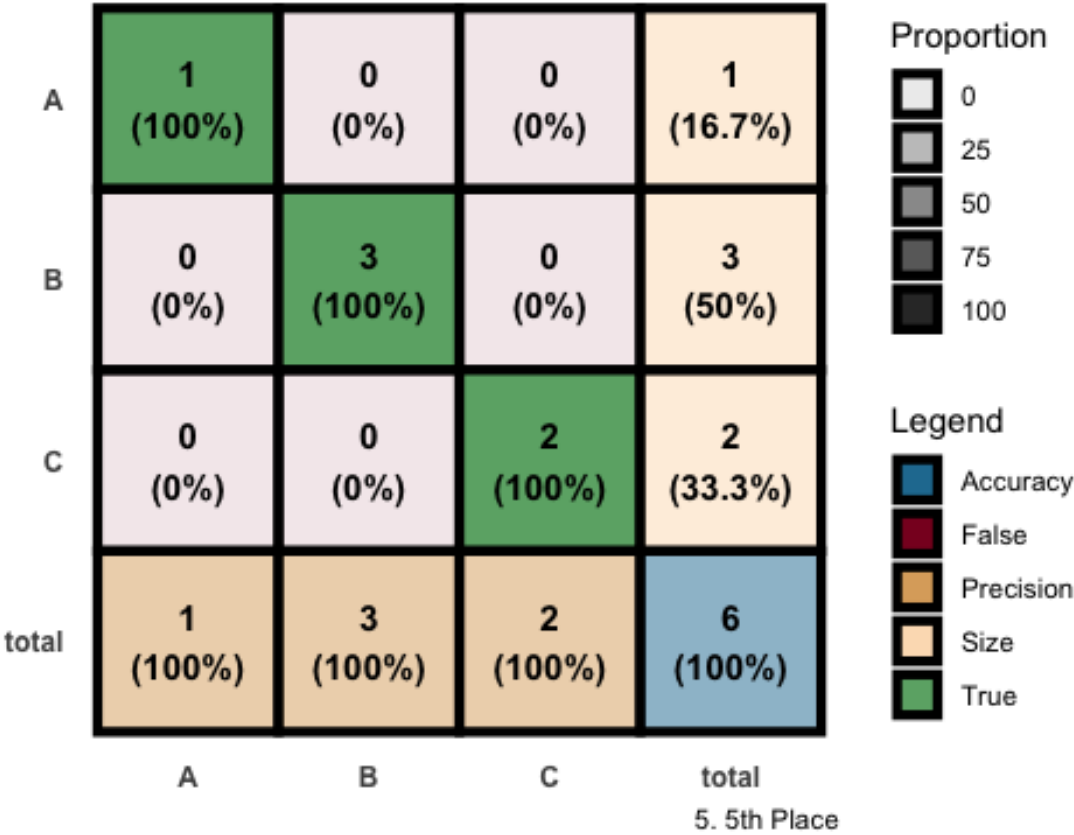
Probability Heatmap



5. 5th Place

External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	18	80	2
2BE	0	51	49
2BF	51	49	0
2BG	2	79	19
2BH	0	0	100
2BI	0	24	76

Exp

B

B

A

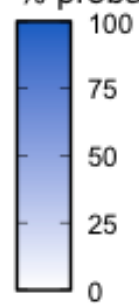
B

C

C

5. 5th Place

% probability



6th Place

Table:

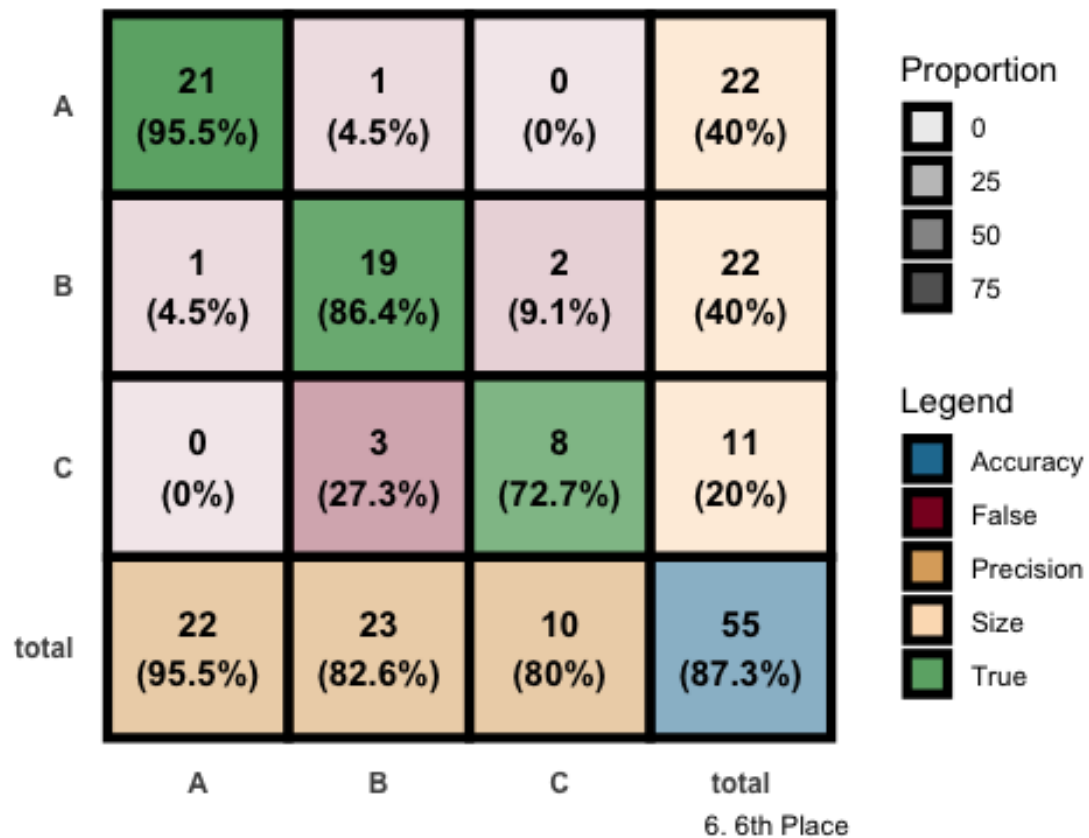
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
87.27%	0.612

Model Coefficients

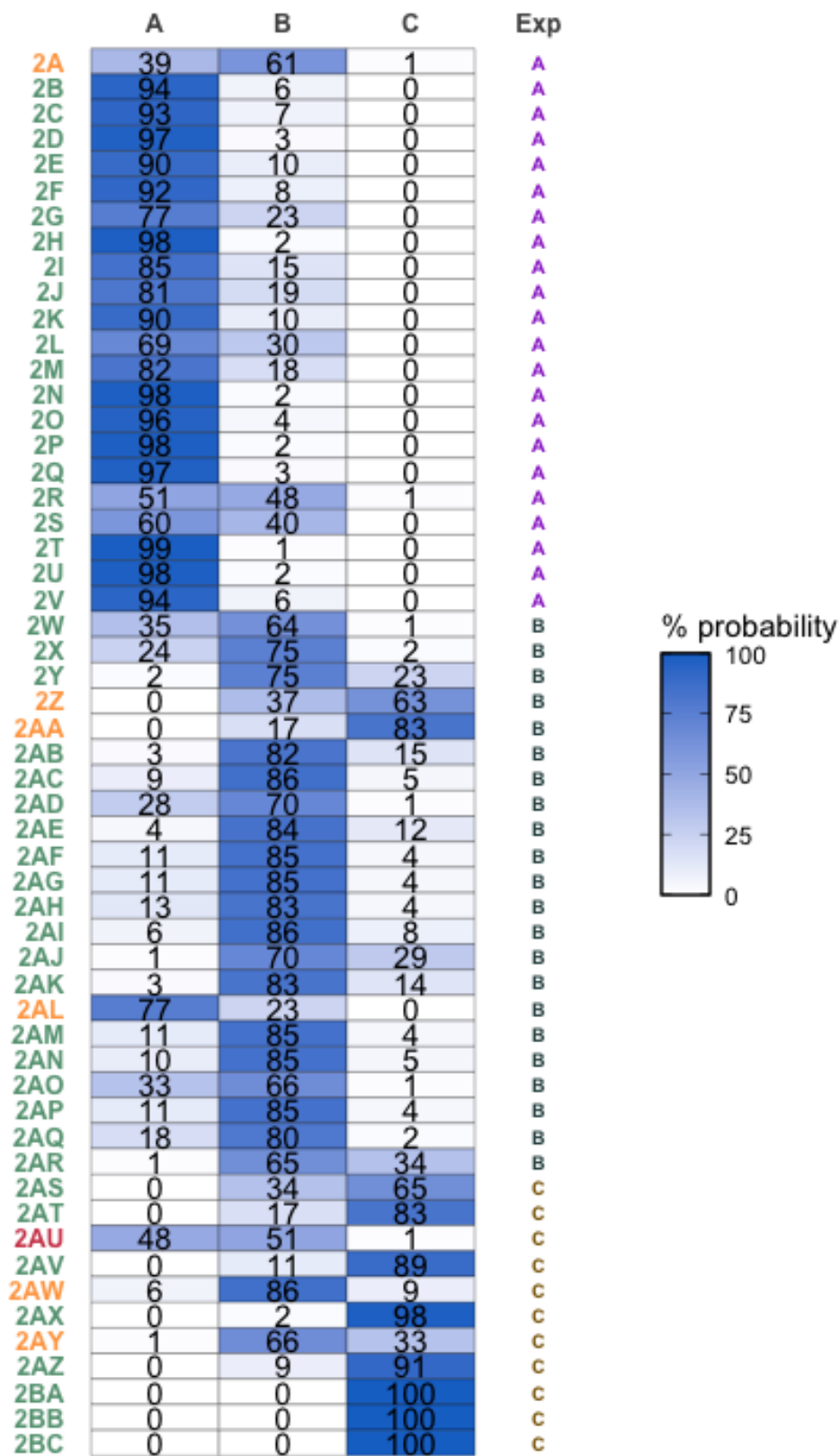
	x
Total	2.0249154
NPA_3	-214.0491915
diff 1 2	-29.6819745
loc.B5	-0.4581857

Training Set Confusion Matrix



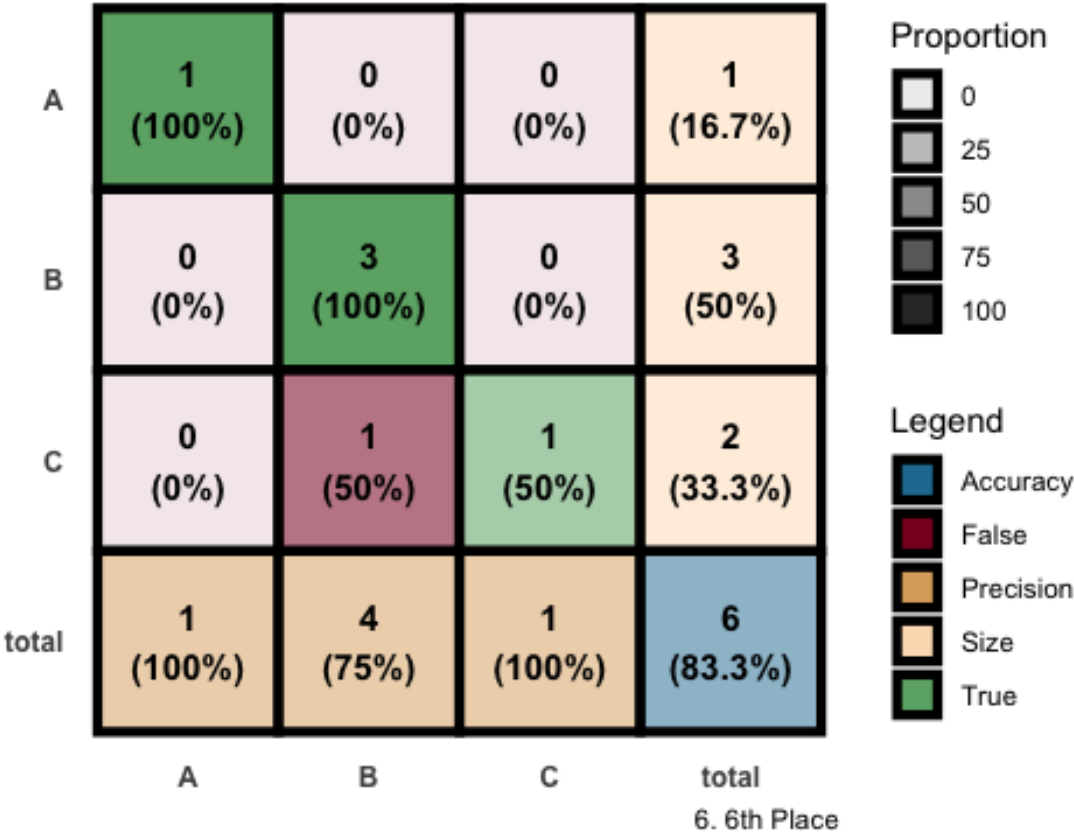
Training Set

Probability Heatmap



External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	16	81	3
2BE	3	80	18
2BF	88	12	0
2BG	2	73	26
2BH	0	0	100
2BI	1	59	41

Exp

B

B

A

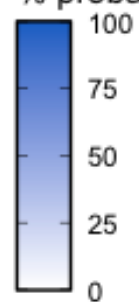
B

C

C

6. 6th Place

% probability



7th Place

Table:

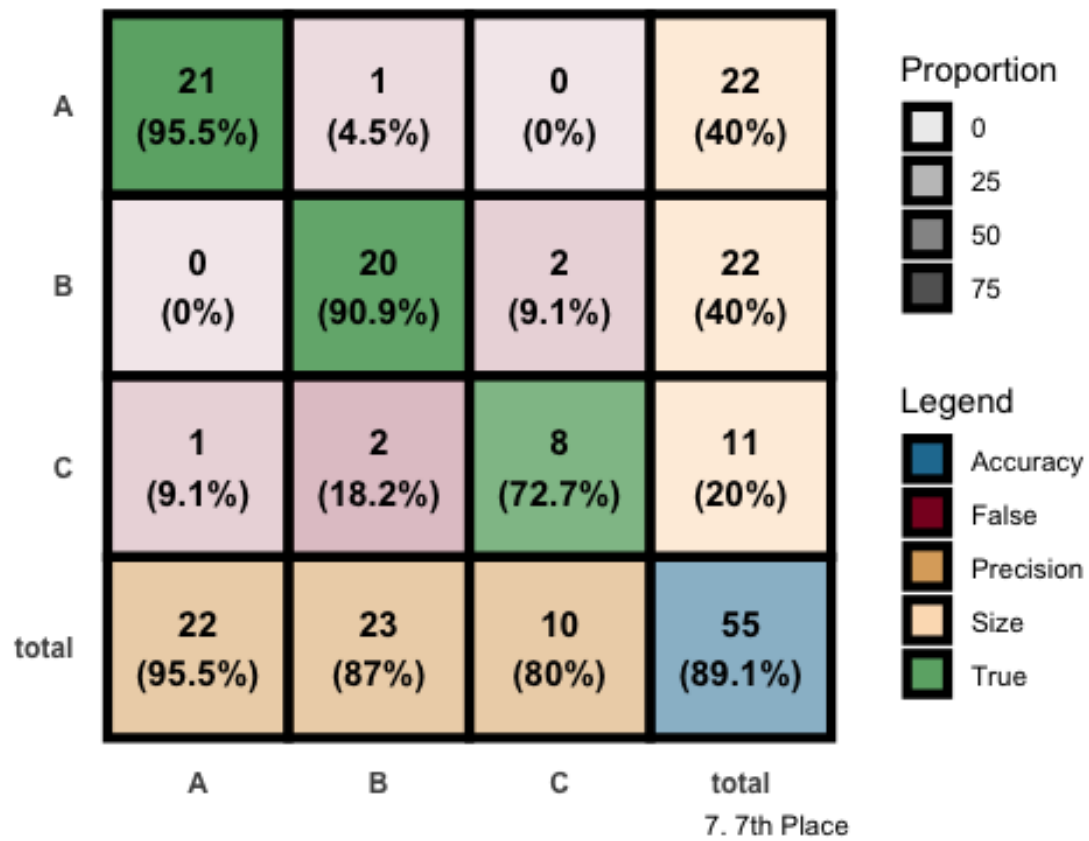
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
89.09%	0.611

Model Coefficients

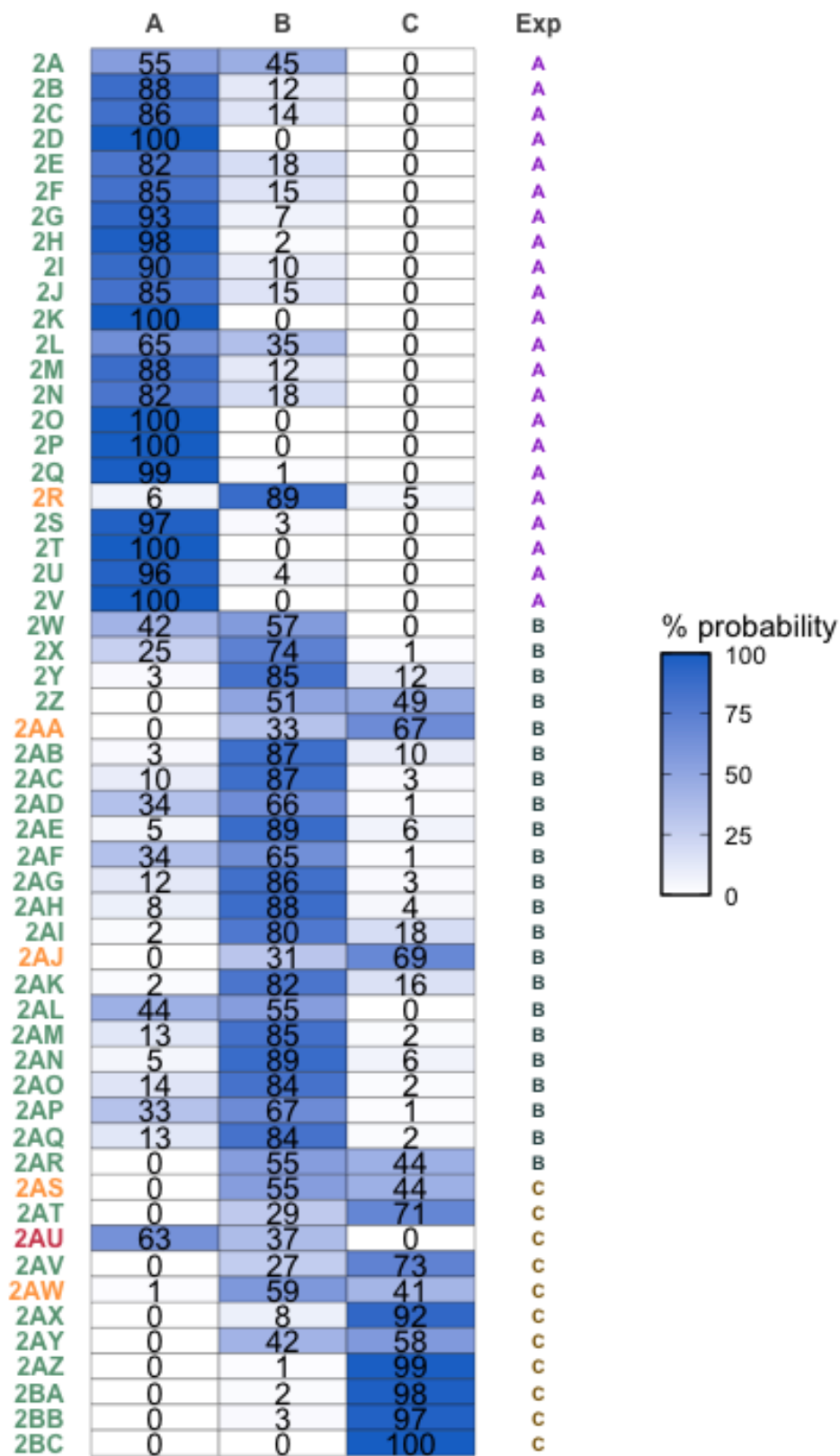
	x
dip_y	-1.445297
Dist(2, 7)	-216.502286
NPA_1	41.710705
NPA_3	-385.833684

Training Set Confusion Matrix



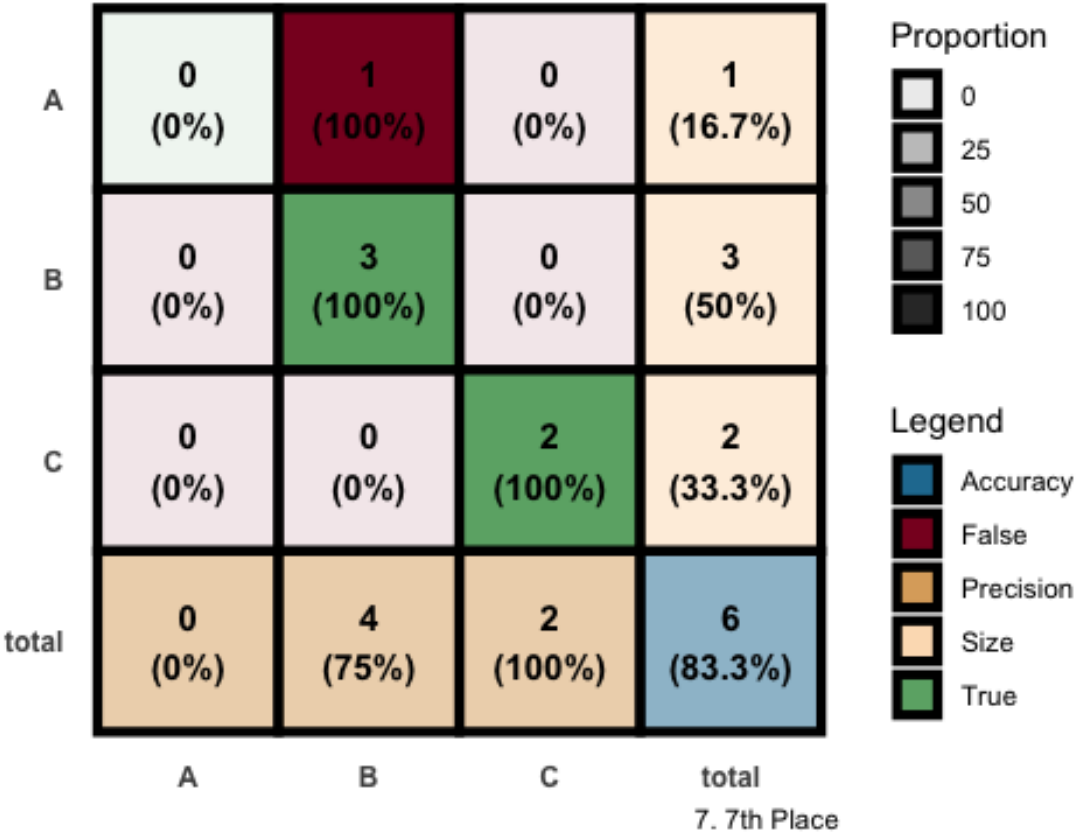
Training Set

Probability Heatmap



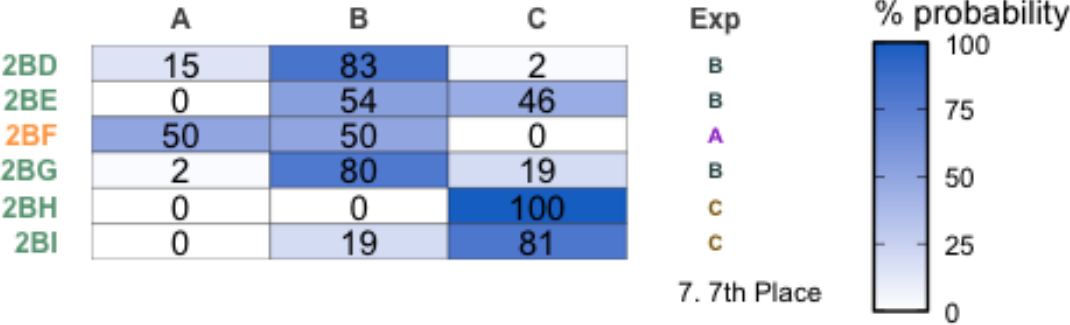
7. 7th Place

External Validation
Confusion Matrix



External Validation

Probability Heatmap



8th Place

Table:

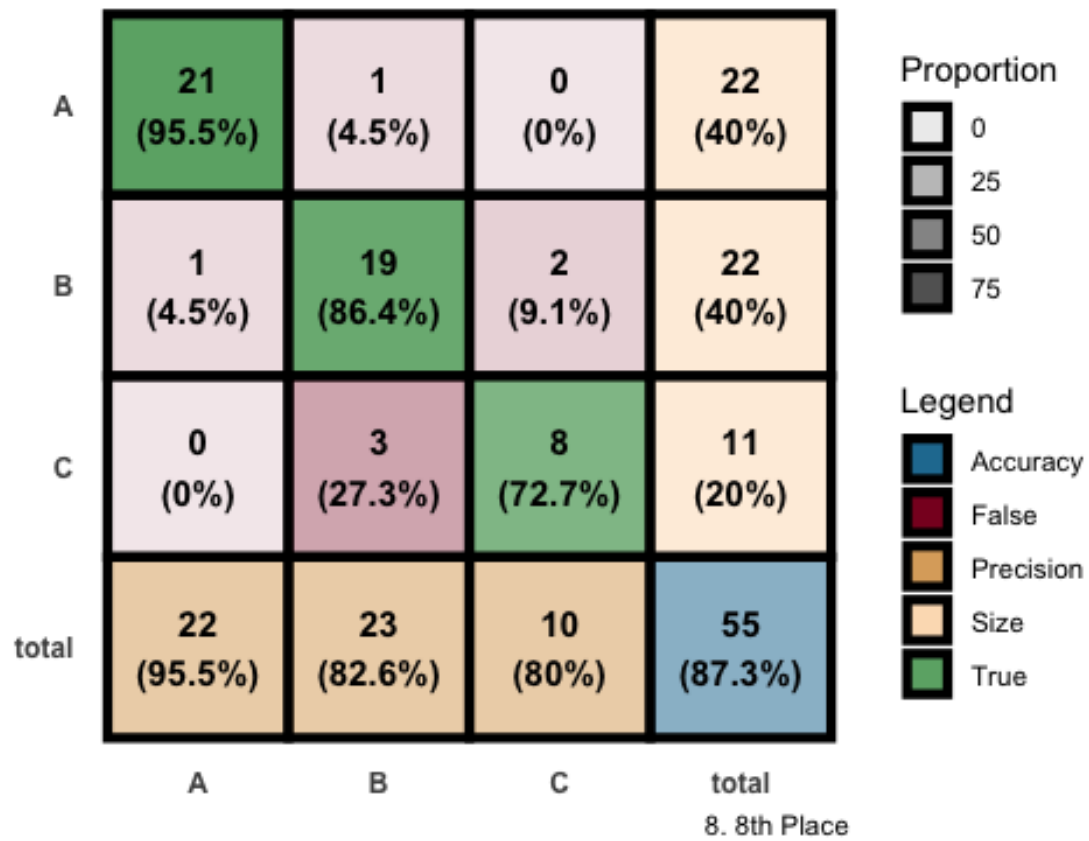
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
87.27%	0.611

Model Coefficients

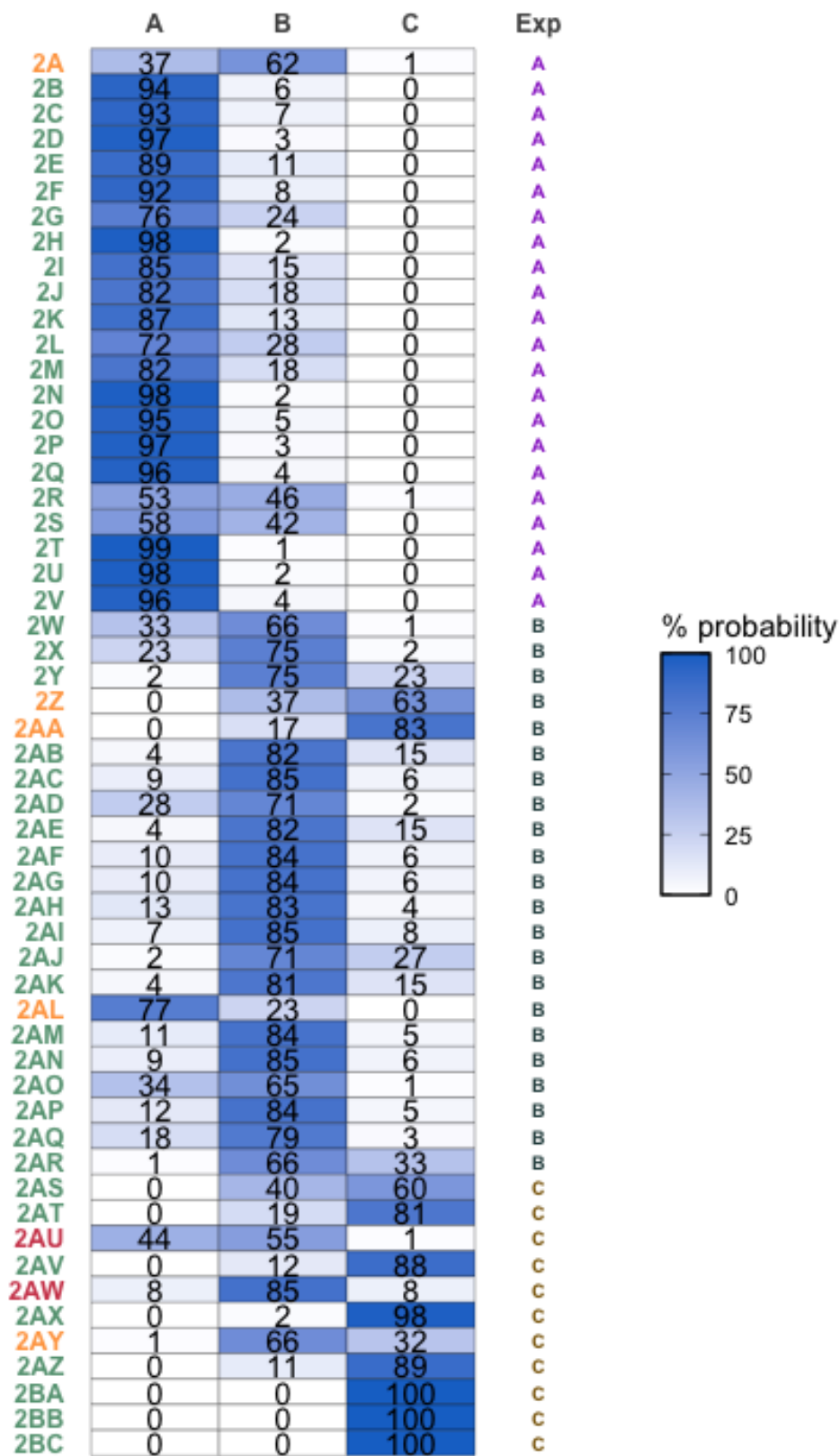
	x
Total	2.103745
NPA_1	28.211591
NPA_3	-185.424470
loc.B5	-0.488480

Training Set Confusion Matrix



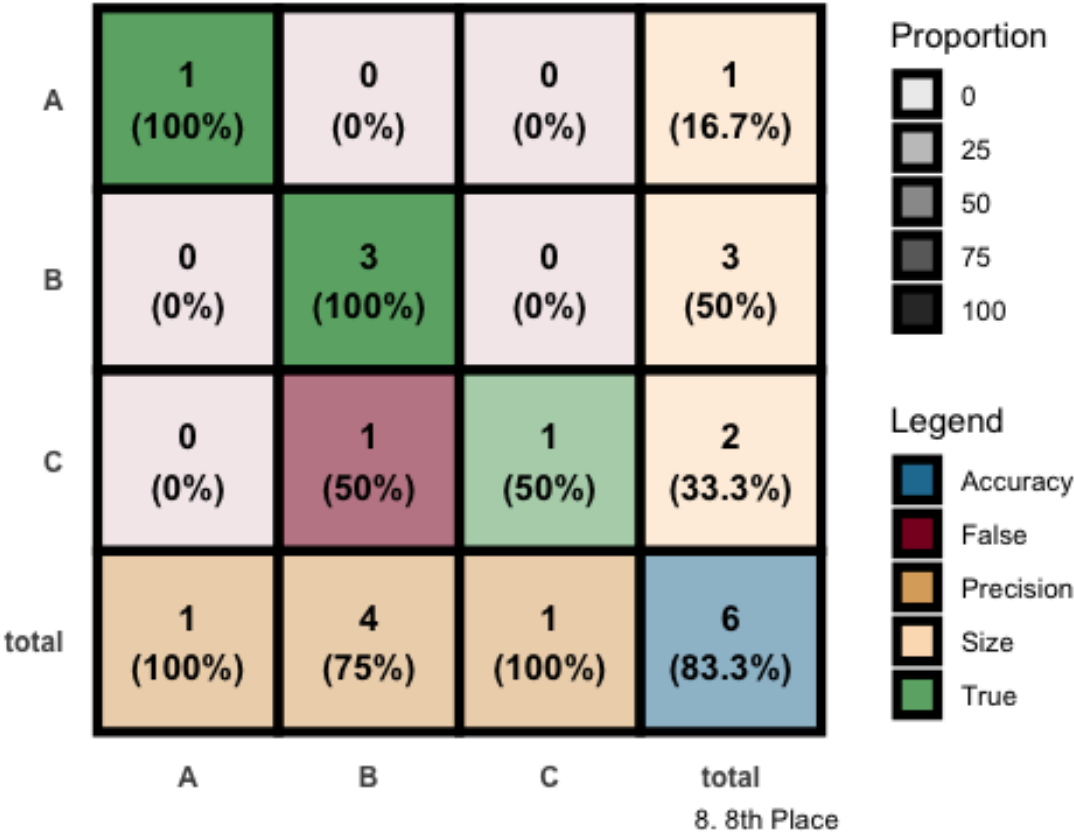
Training Set

Probability Heatmap



External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	16	80	3
2BE	3	80	17
2BF	89	10	0
2BG	2	73	25
2BH	0	0	100
2BI	1	58	41

Exp

B

B

A

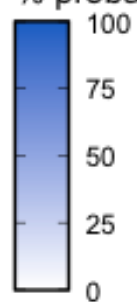
B

C

C

8. 8th Place

% probability



9th Place

Table:

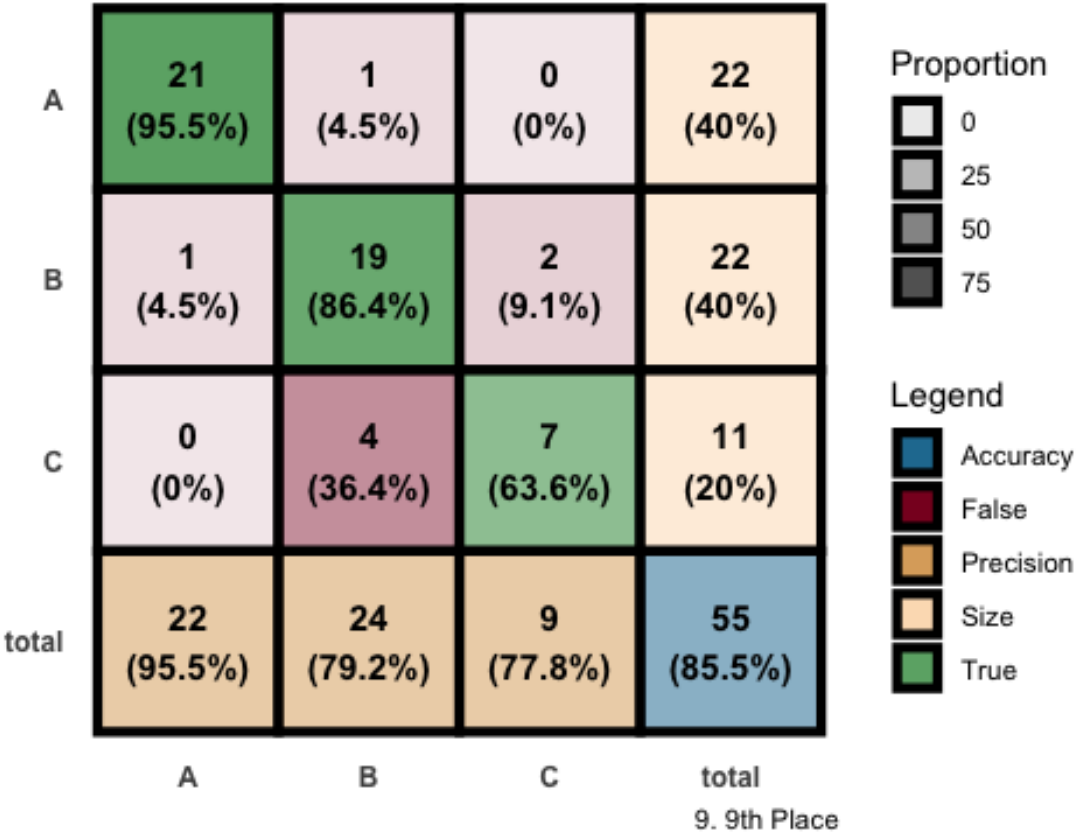
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.45%	0.61

Model Coefficients

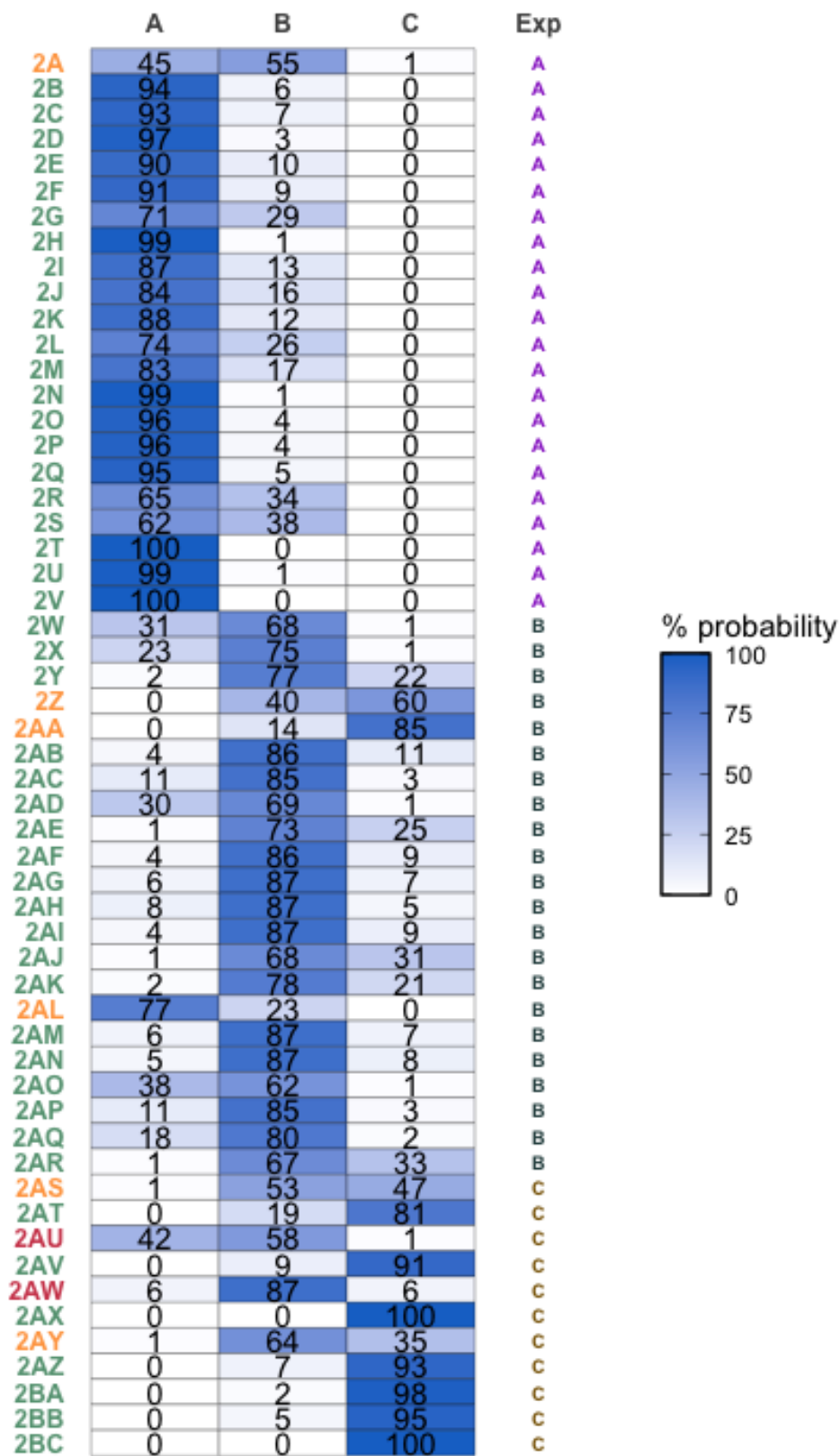
	x
Total	2.5781439
diff 1 2	-35.1062475
diff 2 3	186.6586708
loc.B5	-0.7037206

Training Set
Confusion Matrix

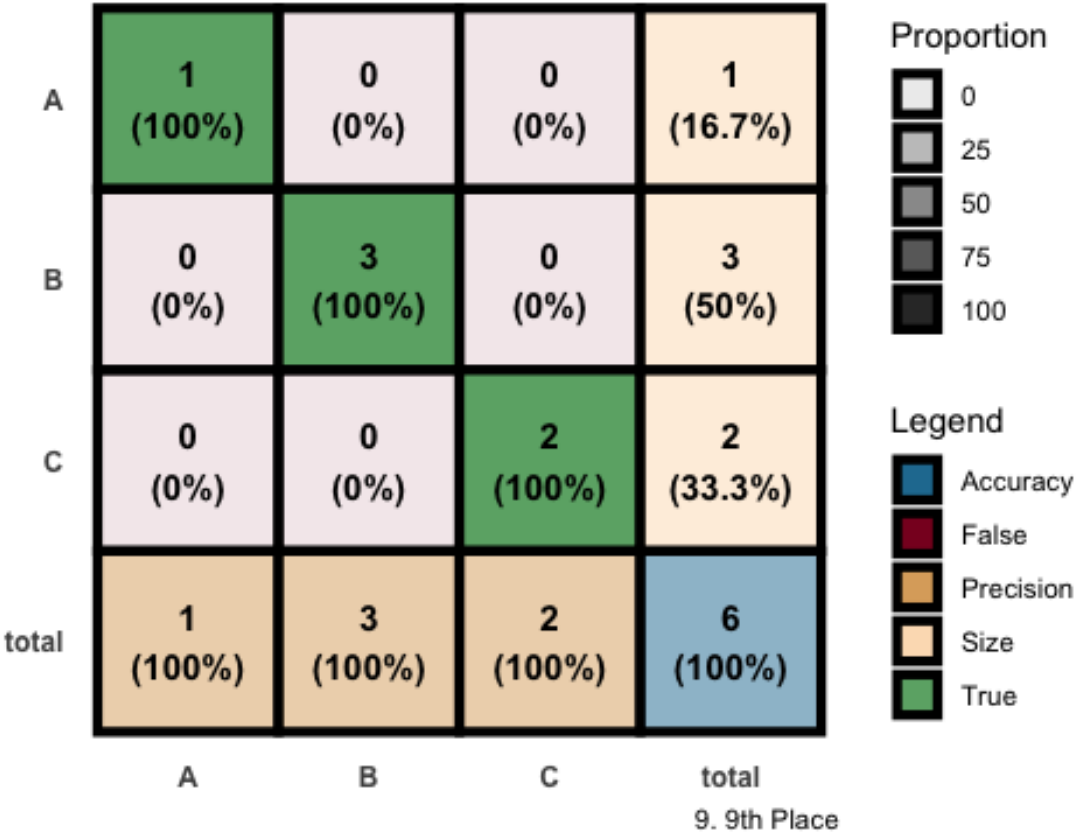


Training Set

Probability Heatmap



External Validation
Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	10	86	4
2BE	5	87	8
2BF	93	7	0
2BG	1	76	23
2BH	0	0	100
2BI	0	46	54

Exp

B

B

A

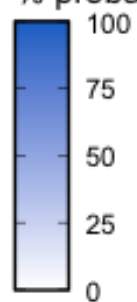
B

C

C

9. 9th Place

% probability



10th Place

Table:

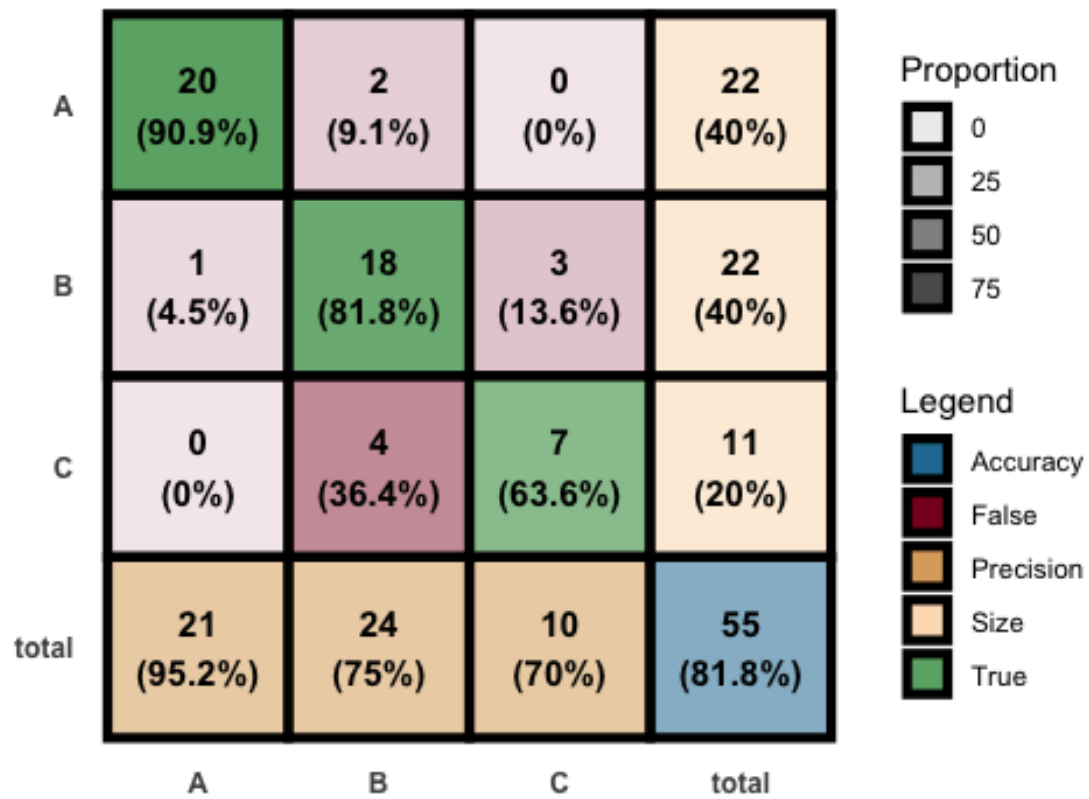
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
81.82%	0.607

Model Coefficients

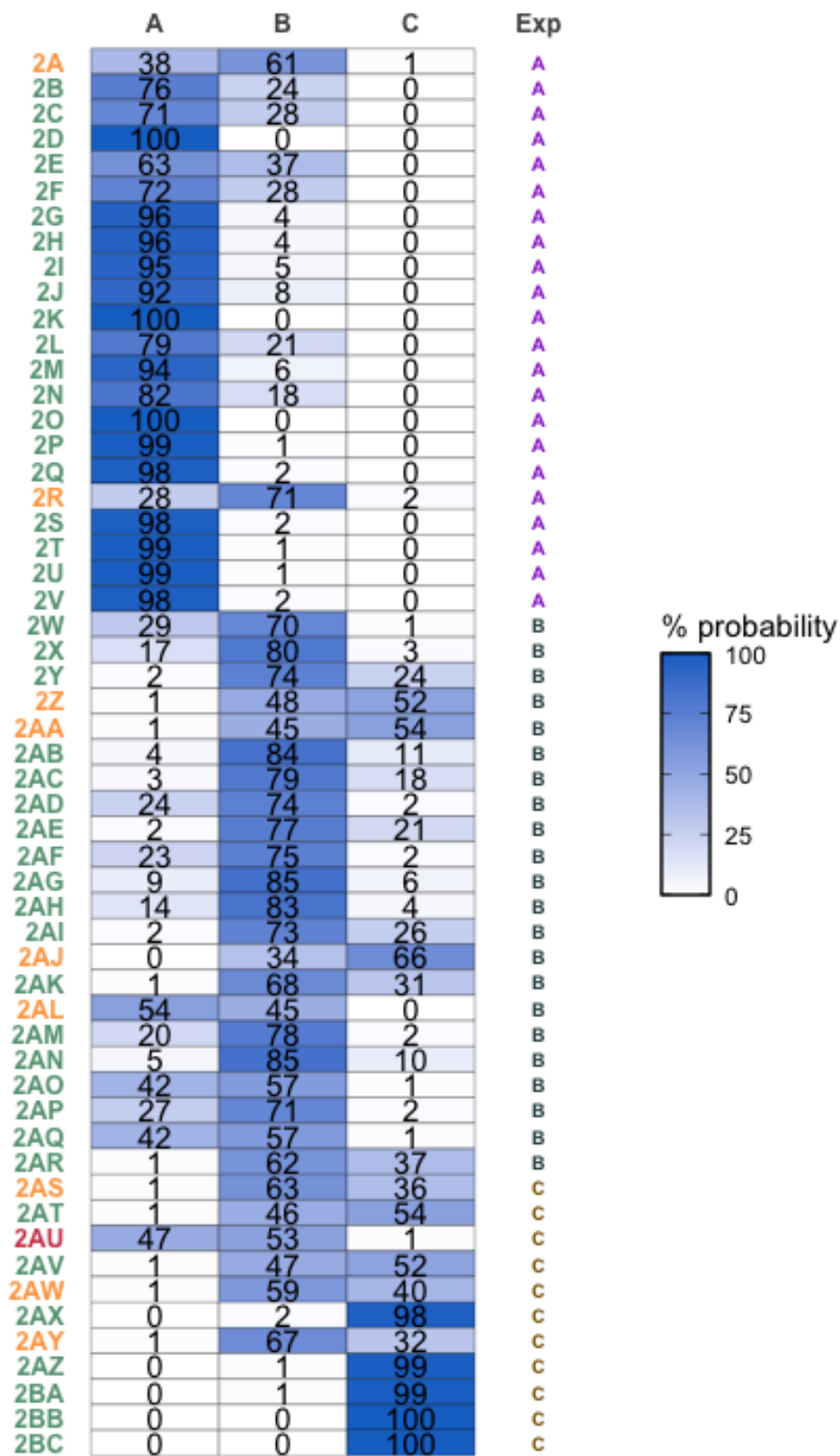
	x
dip_y	-2.2426235
NPA_3	-259.8157729
diff 1 2	-40.3407156
L	-0.4205848

Training Set
Confusion Matrix

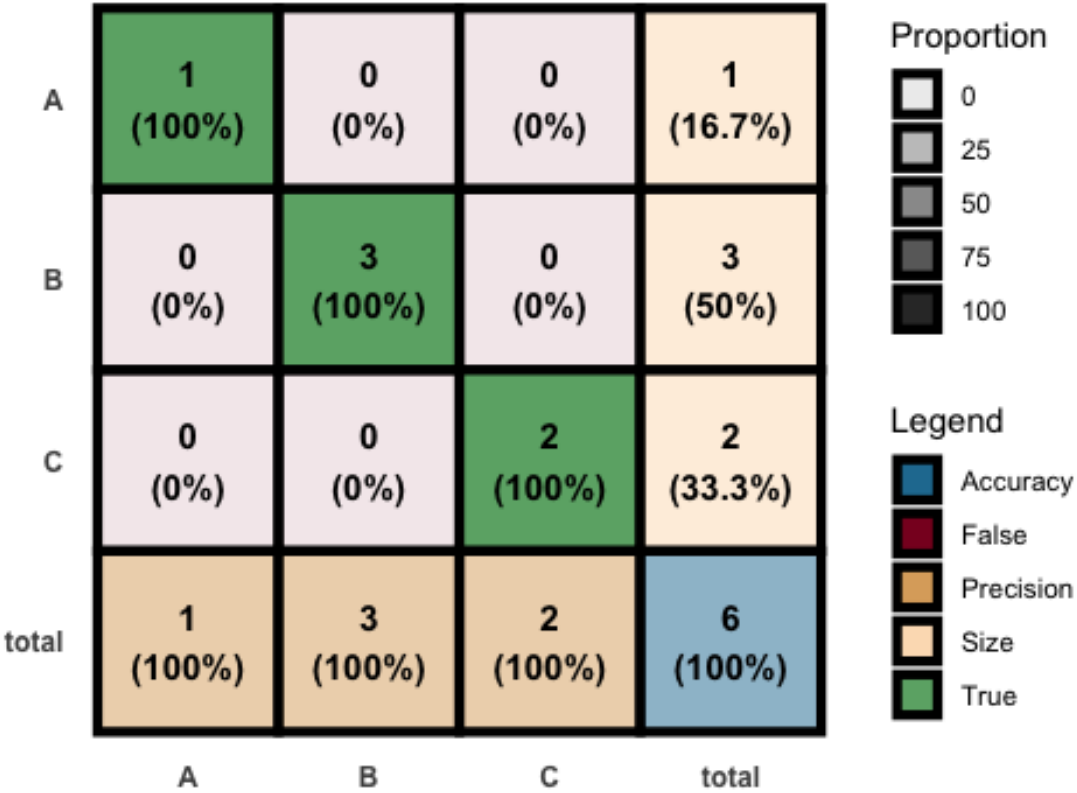


Training Set

Probability Heatmap



External Validation
Confusion Matrix



External Validation

Probability Heatmap

