

Aldehyde Deuteration - Model Search – Non-ordinal

Table of Contents

Model Search.....	1
Section Results Summary	1
1st Place.....	2
2nd Place.....	9
3rd Place.....	16
4th Place.....	23
5th Place.....	30
6th Place.....	37
7th Place.....	44
8th Place.....	51
9th Place.....	58
10th Place.....	65

Model Search

formula	McFadden R2
class ~ -2-3- + dip_x + dip_y + NPA_1	0.716
class ~ -2-3- + dip_x + dip_y + diff 1 2	0.707
class ~ dip_x + dip_y + Dist(2, 3) + NPA_1	0.699
class ~ dip_x + dip_y + Dist(2, 3) + diff 1 2	0.699
class ~ dip_x + dip_y + NPA_1 + NPA_3	0.699
class ~ dip_x + dip_y + NPA_3 + diff 1 2	0.690
class ~ Total + NPA_1 + NPA_3 + loc.B5	0.687
class ~ Total + NPA_3 + diff 1 2 + loc.B5	0.678
class ~ Total + diff 1 2 + diff 2 3 + loc.B5	0.672
class ~ dip_x + dip_y + NPA_1 + diff 1 2	0.671

Section Results Summary

	Training Accuracy	Test Accuracy
1. 1st Place	85.71	92.3
2. 2nd Place	83.33	84.6
3. 3rd Place	88.10	92.3
4. 4th Place	85.71	84.6
5. 5th Place	92.86	84.6

	Training Accuracy	Test Accuracy
6. 6th Place	92.86	84.6
7. 7th Place	85.71	92.3
8. 8th Place	85.71	92.3
9. 9th Place	88.10	92.3
10. 10th Place	85.71	84.6

1st Place

Table:

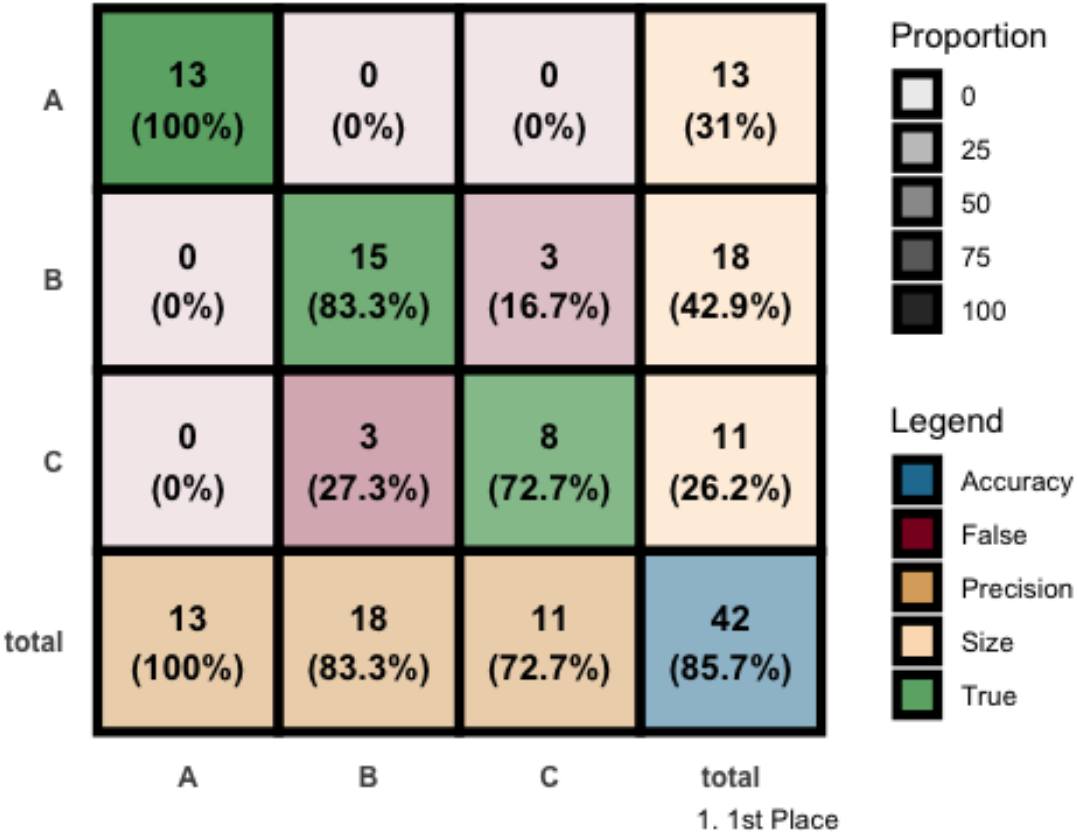
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.716

Model Coefficients

	(Intercept)	-2-3-	dip_x	dip_y	NPA_1
2	-231.4548	0.1184213	-11.96237	-17.85471	208.7282
3	255.8387	-0.1478514	-11.63987	-18.64964	221.9646

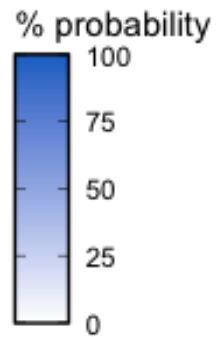
Training Set
Confusion Matrix



Training Set

Probability Heatmap

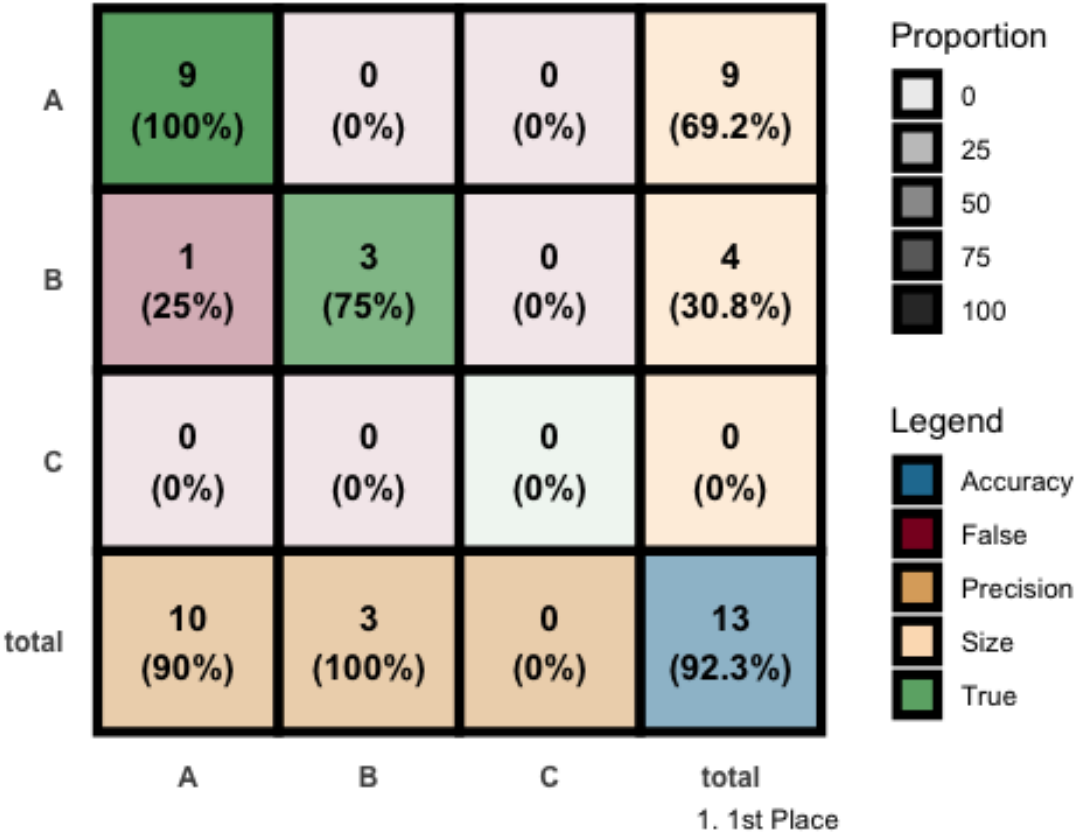
	A	B	C	Exp
2A	48	44	8	A
2G	100	0	0	A
2H	100	0	0	A
2L	100	0	0	A
2M	100	0	0	A
2O	100	0	0	A
2R	92	7	1	A
2S	93	7	0	A
2T	100	0	0	A
2U	100	0	0	A
2V	100	0	0	A
2J	100	0	0	A
2P	100	0	0	A
2Y	0	81	19	B
2Z	0	44	56	B
2AA	0	37	63	B
2AE	0	96	4	B
2AF	0	99	1	B
2AG	0	98	2	B
2AH	0	96	4	B
2AJ	0	68	32	B
2AK	0	92	8	B
2AM	0	87	13	B
2AO	0	99	1	B
2W	8	70	22	B
2X	0	71	29	B
2AB	0	74	26	B
2AC	0	70	30	B
2AP	0	88	12	B
2AQ	0	99	1	B
2AR	0	39	61	B
2AS	0	56	44	C
2AT	0	31	69	C
2AU	41	48	11	C
2AV	0	34	66	C
2AW	18	64	18	C
2AX	0	35	65	C
2AY	0	32	68	C
2AZ	0	17	83	C
2BA	0	0	100	C
2BB	0	17	83	C
2BC	0	1	99	C



1. 1st Place

Test Set

Confusion Matrix



Test Set

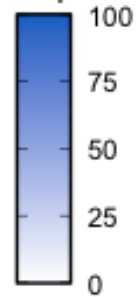
Probability Heatmap

	A	B	C
2B	100	0	0
2C	100	0	0
2D	100	0	0
2E	100	0	0
2F	100	0	0
2I	100	0	0
2K	100	0	0
2N	100	0	0
2Q	100	0	0
2AD	92	6	2
2AI	0	91	9
2AL	43	56	1
2AN	0	93	7

Exp

A
A
A
A
A
A
A
A
A
A
B
B
B
B

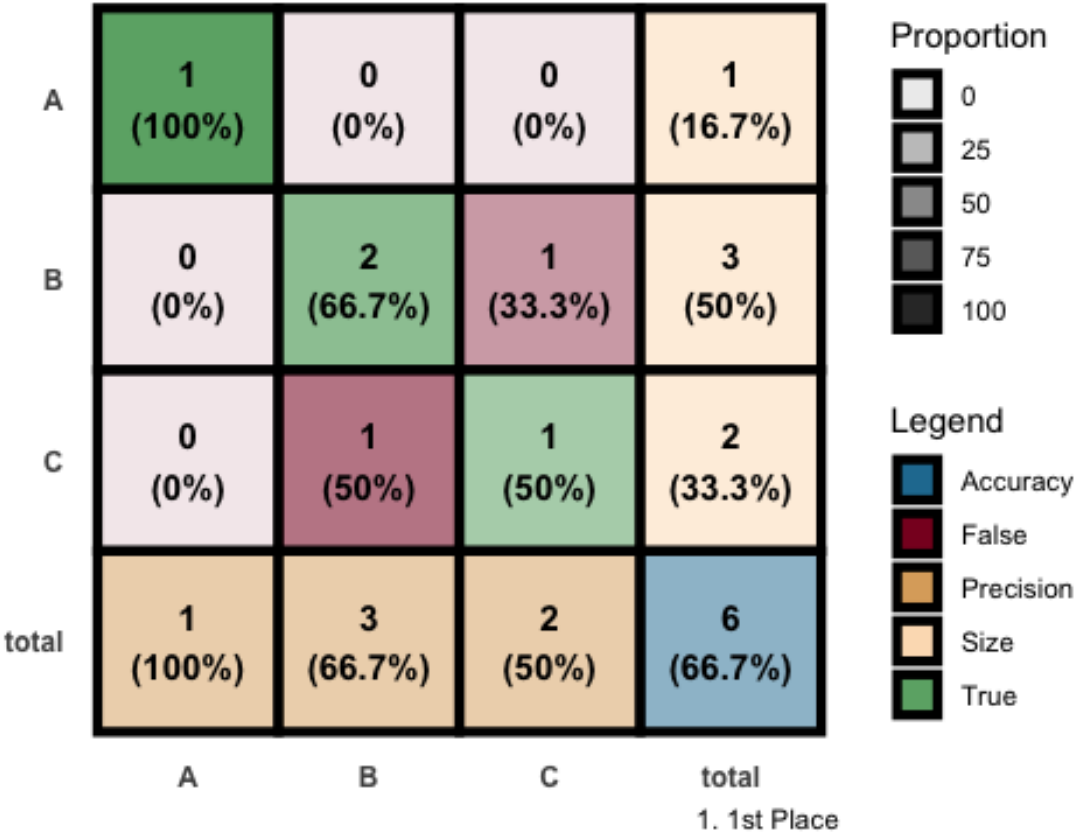
% probability



1. 1st Place

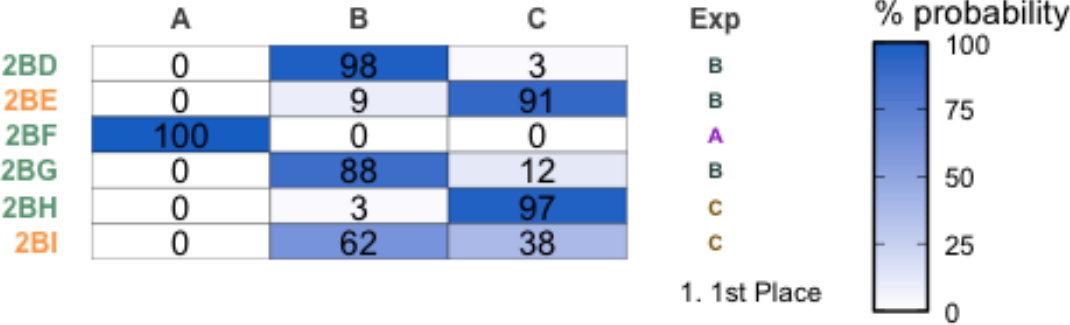
External Validation

Confusion Matrix



External Validation

Probability Heatmap



2nd Place

Table:

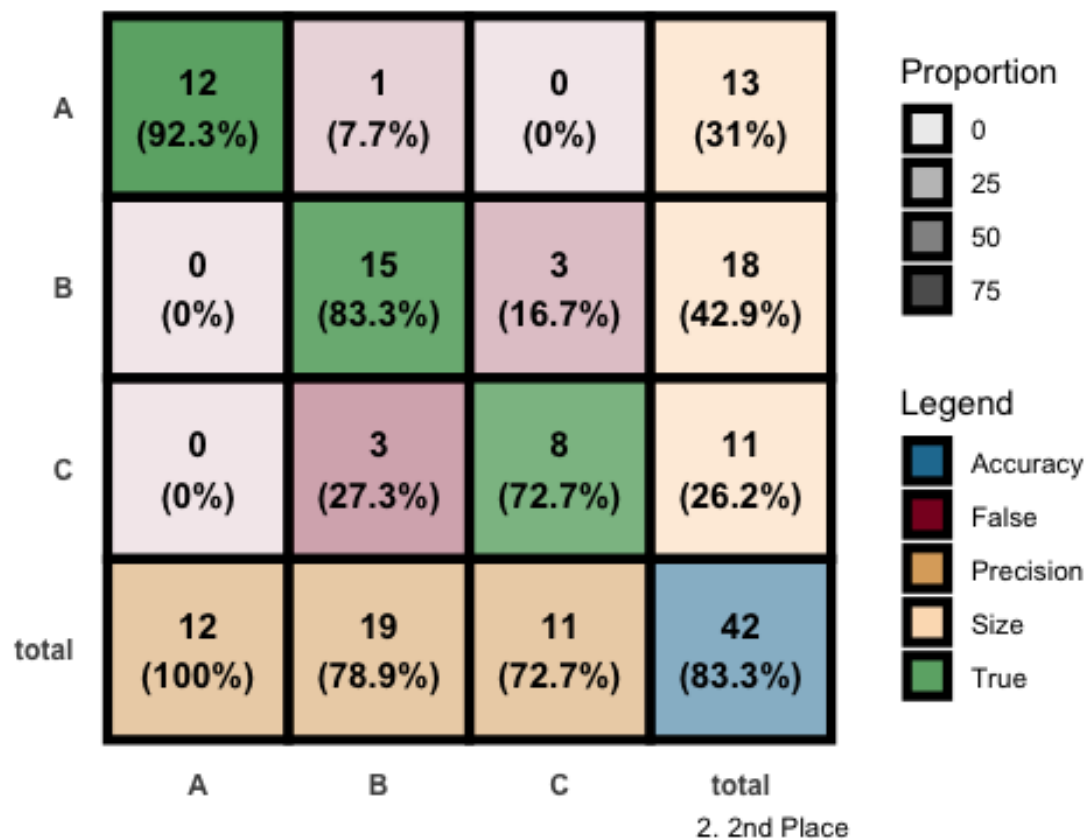
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
83.33%	0.707

Model Coefficients

	(Intercept)	-2-3-	dip_x	dip_y	diff 1 2
2	-285.5935	0.1937231	-11.55657	-17.63469	-201.2935
3	327.8016	-0.1380863	-11.03687	-18.45009	-215.1336

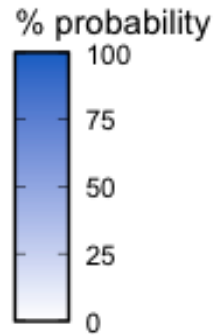
Training Set Confusion Matrix



Training Set

Probability Heatmap

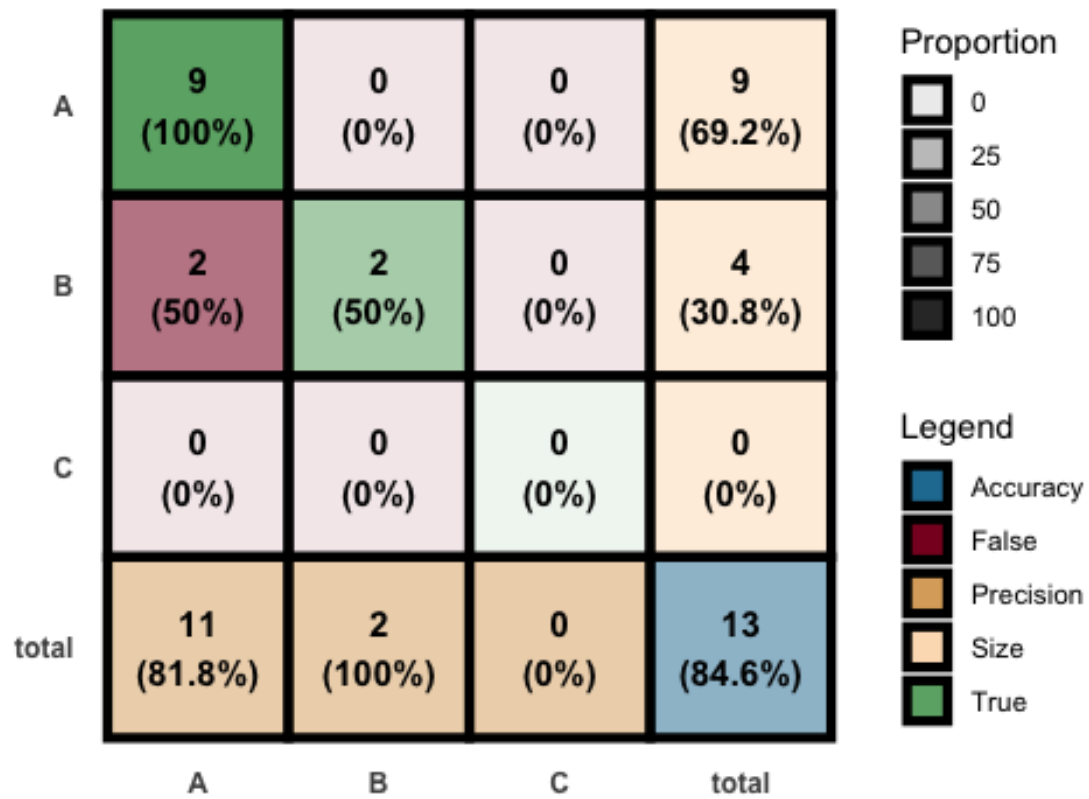
	A	B	C	Exp
2A	35	55	10	A
2G	100	0	0	A
2H	100	0	0	A
2L	100	0	0	A
2M	100	0	0	A
2O	100	0	0	A
2R	91	8	1	A
2S	92	8	0	A
2T	100	0	0	A
2U	100	0	0	A
2V	100	0	0	A
2J	100	0	0	A
2P	100	0	0	A
2Y	0	84	16	B
2Z	0	42	58	B
2AA	0	34	66	B
2AE	0	98	2	B
2AF	0	100	0	B
2AG	0	100	0	B
2AH	0	98	2	B
2AJ	0	69	31	B
2AK	0	95	5	B
2AM	0	90	10	B
2AO	0	100	0	B
2W	7	70	23	B
2X	0	69	31	B
2AB	0	76	24	B
2AC	0	69	31	B
2AP	0	91	9	B
2AQ	0	100	0	B
2AR	0	35	65	B
2AS	0	53	47	C
2AT	0	27	73	C
2AU	41	48	10	C
2AV	0	30	70	C
2AW	34	52	14	C
2AX	0	43	57	C
2AY	0	27	73	C
2AZ	0	14	86	C
2BA	0	0	100	C
2BB	0	13	87	C
2BC	0	0	100	C



2. 2nd Place

Test Set

Confusion Matrix



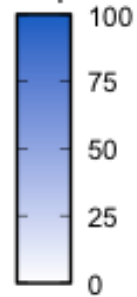
2. 2nd Place

Test Set

Probability Heatmap

	A	B	C	Exp
2B	100	0	0	A
2C	100	0	0	A
2D	100	0	0	A
2E	100	0	0	A
2F	100	0	0	A
2I	100	0	0	A
2K	100	0	0	A
2N	100	0	0	A
2Q	100	0	0	A
2AD	87	9	4	B
2AI	0	95	5	B
2AL	53	46	0	B
2AN	0	97	3	B

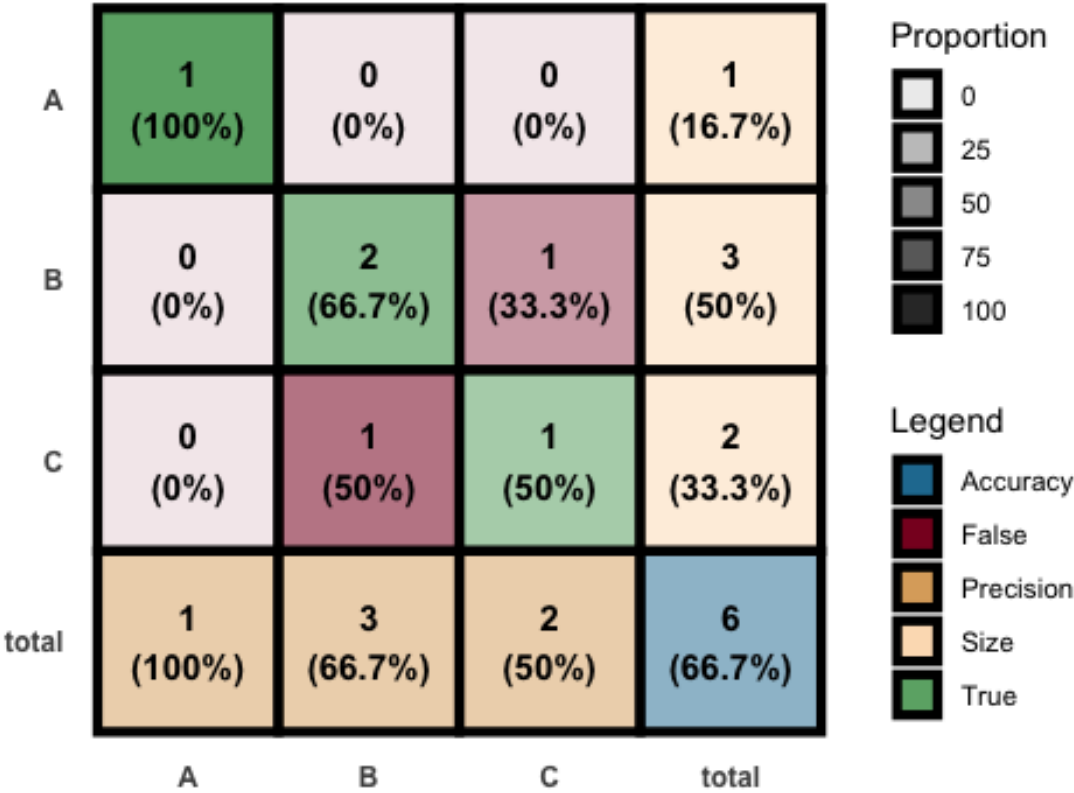
% probability



2. 2nd Place

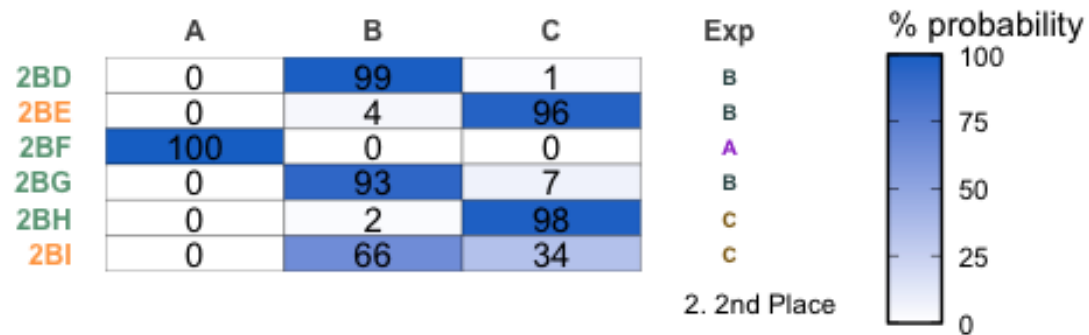
External Validation

Confusion Matrix



External Validation

Probability Heatmap



3rd Place

Table:

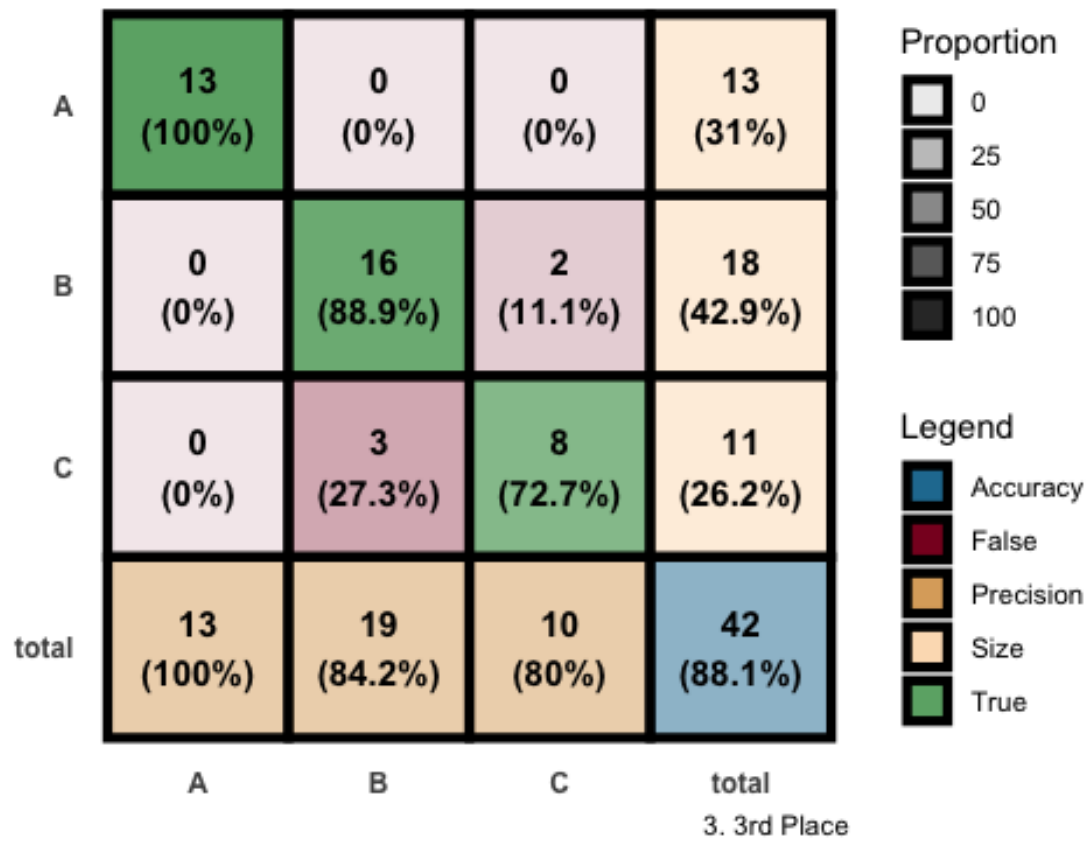
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
88.1%	0.699

Model Coefficients

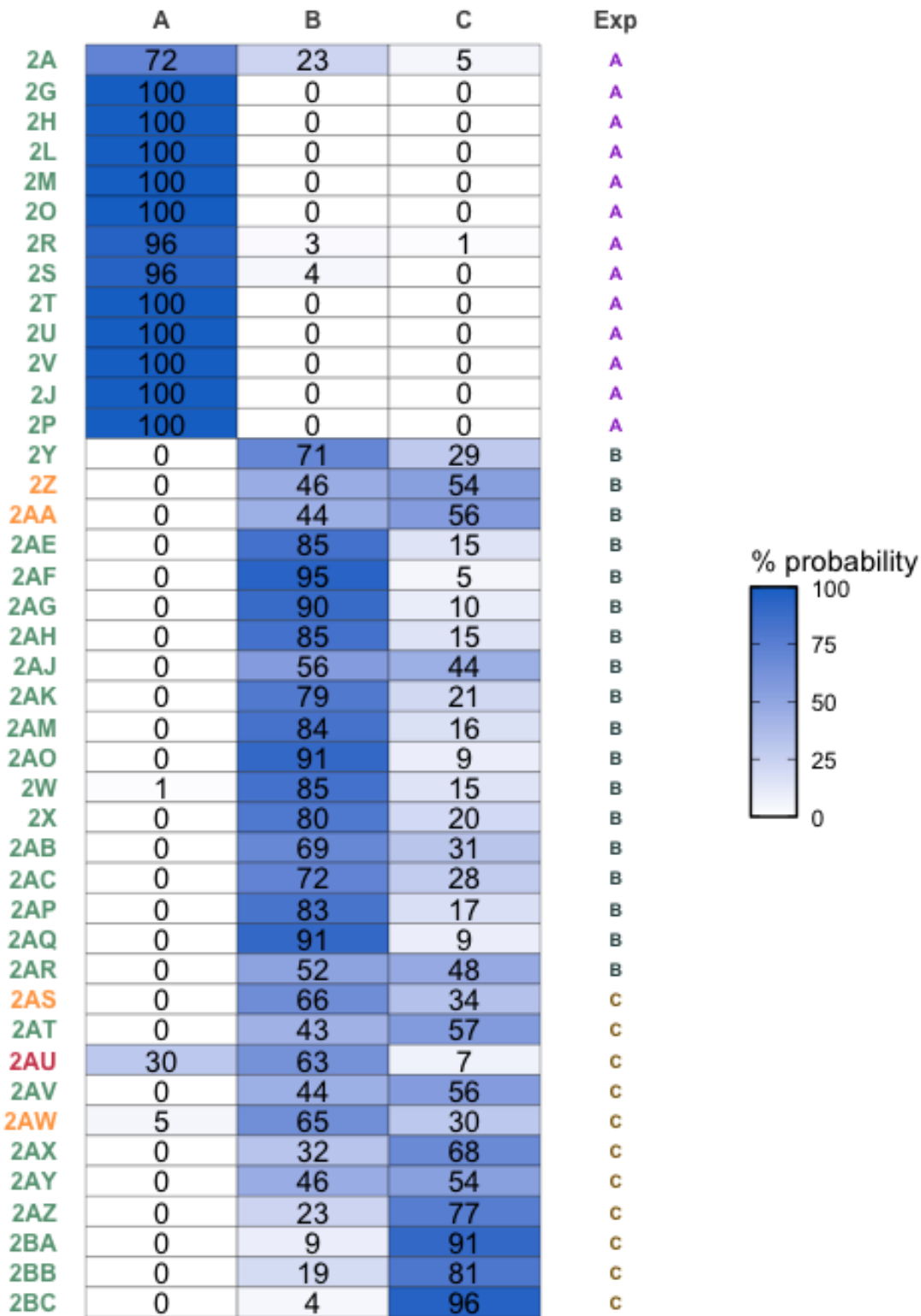
	(Intercept)	dip_x	dip_y	Dist(2, 3)	NPA_1
2	462.3908	-22.00662	-33.83547	-400.2782	430.6610
3	-465.2533	-21.84399	-34.51954	370.3708	442.1676

Training Set Confusion Matrix



Training Set

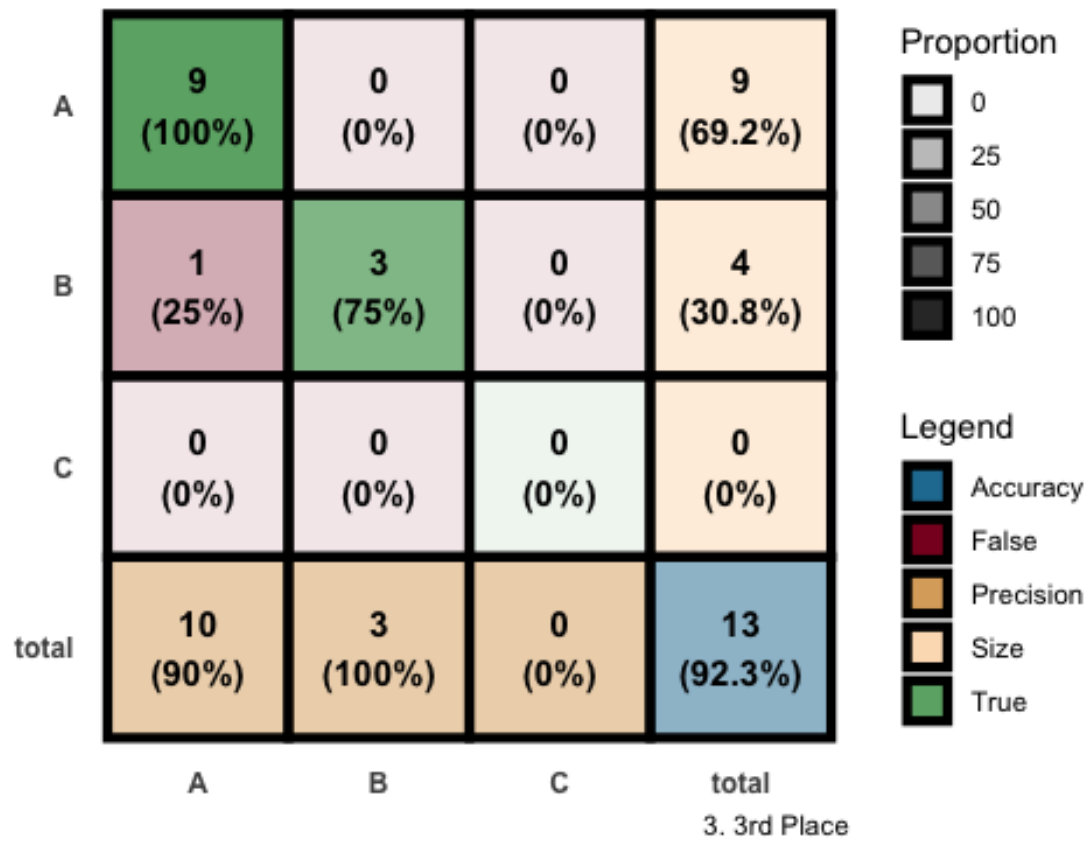
Probability Heatmap



3. 3rd Place

Test Set

Confusion Matrix



Test Set

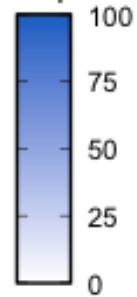
Probability Heatmap

	A	B	C
2B	100	0	0
2C	100	0	0
2D	100	0	0
2E	100	0	0
2F	100	0	0
2I	100	0	0
2K	100	0	0
2N	100	0	0
2Q	100	0	0
2AD	99	1	0
2AI	0	76	24
2AL	46	50	4
2AN	0	79	21

Exp

A
A
A
A
A
A
A
A
A
A
B
B
B
B

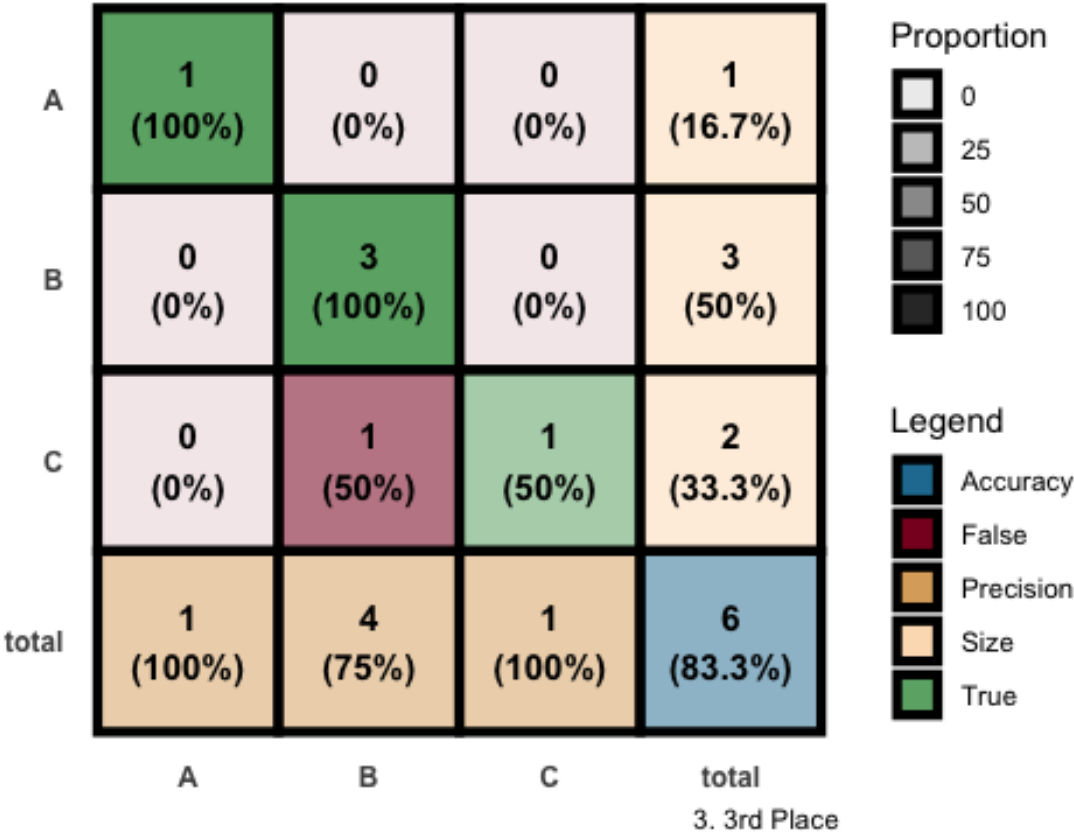
% probability



3. 3rd Place

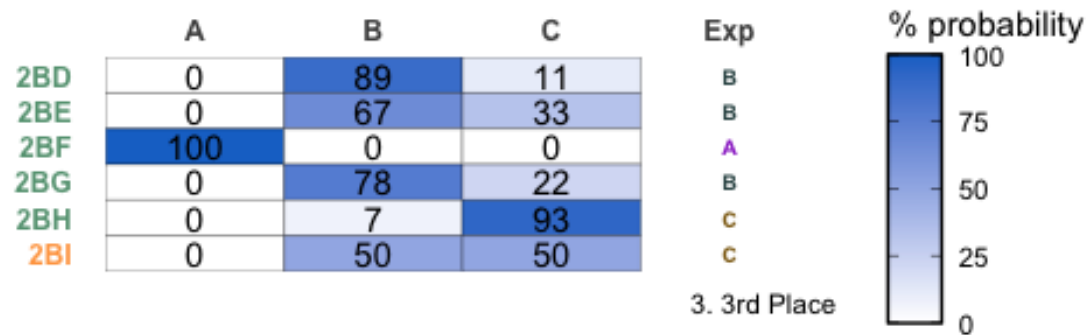
External Validation

Confusion Matrix



External Validation

Probability Heatmap



4th Place

Table:

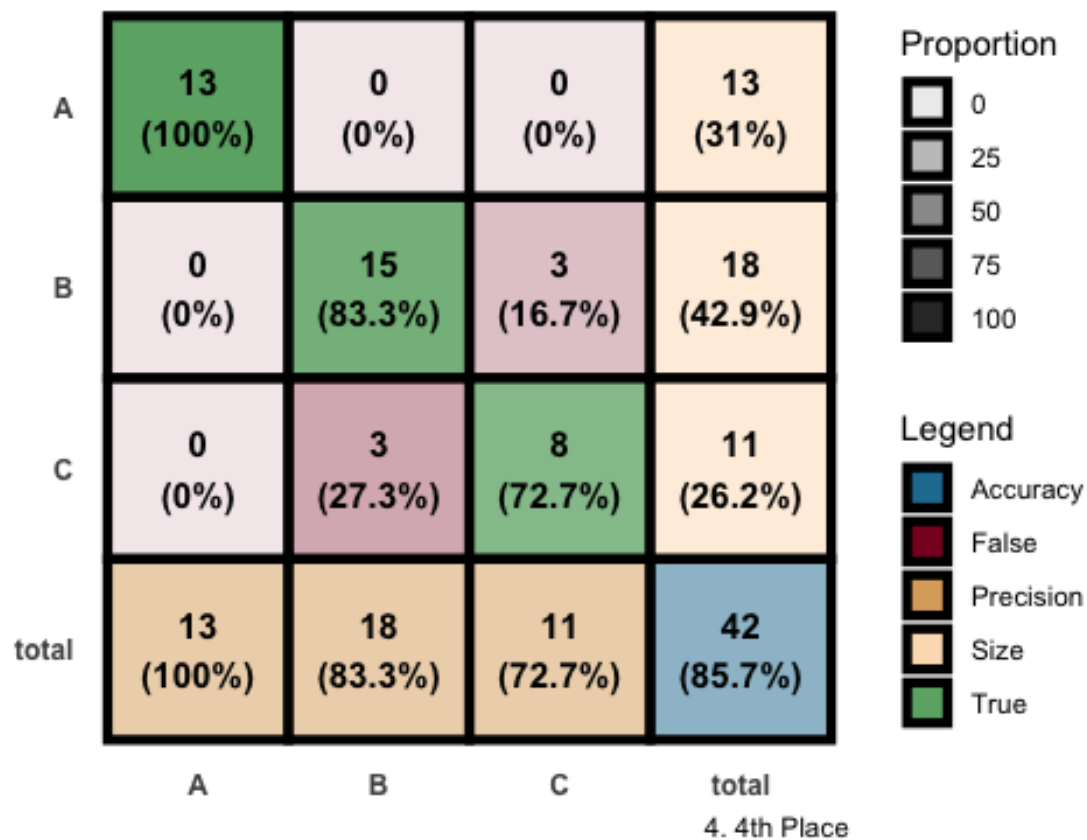
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.699

Model Coefficients

	(Intercept)	dip_x	dip_y	Dist(2, 3)	diff 1 2
2	661.7346	-22.71409	-36.60076	-398.7258	-475.5286
3	-506.7224	-22.42204	-37.26460	576.4101	-487.2459

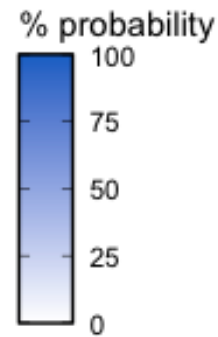
Training Set Confusion Matrix



Training Set

Probability Heatmap

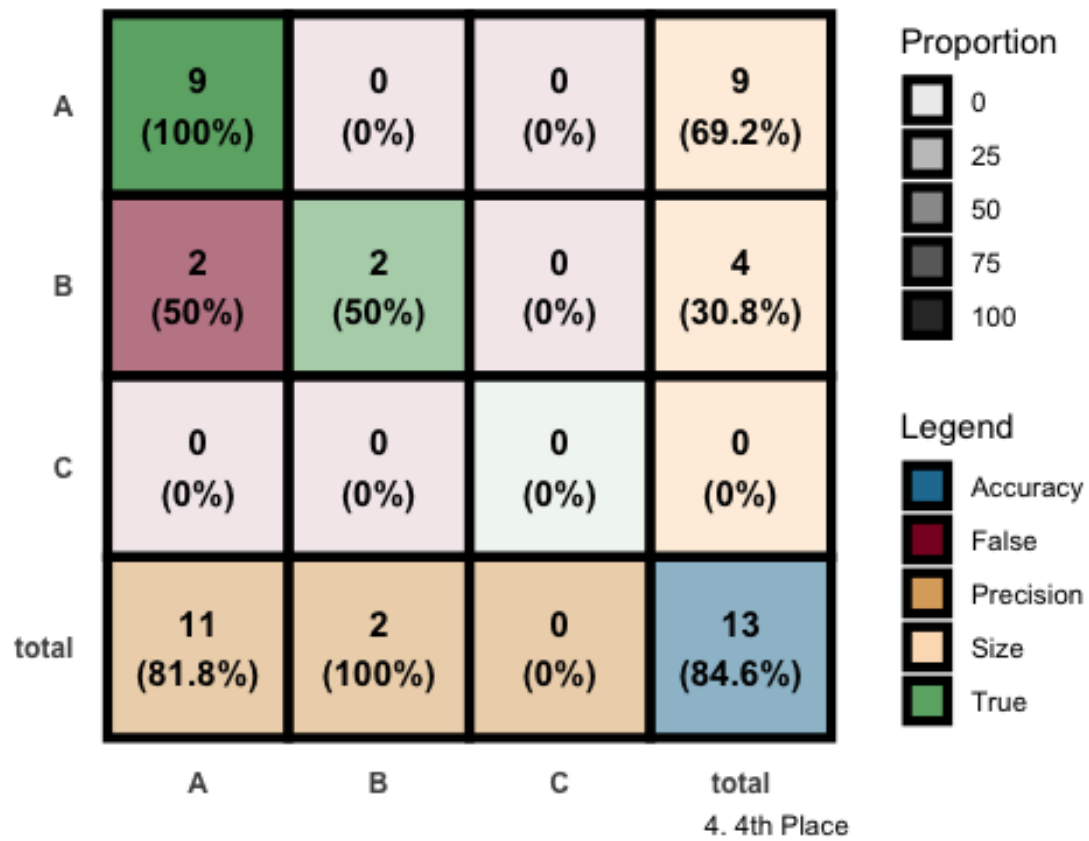
	A	B	C	Exp
2A	60	33	7	A
2G	100	0	0	A
2H	100	0	0	A
2L	100	0	0	A
2M	100	0	0	A
2O	100	0	0	A
2R	94	4	2	A
2S	95	5	0	A
2T	100	0	0	A
2U	100	0	0	A
2V	100	0	0	A
2J	100	0	0	A
2P	100	0	0	A
2Y	0	72	28	B
2Z	0	43	57	B
2AA	0	41	59	B
2AE	0	90	10	B
2AF	0	98	2	B
2AG	0	95	5	B
2AH	0	89	11	B
2AJ	0	54	46	B
2AK	0	83	17	B
2AM	0	87	13	B
2AO	0	95	5	B
2W	0	86	13	B
2X	0	80	20	B
2AB	0	68	32	B
2AC	0	73	27	B
2AP	0	84	16	B
2AQ	0	95	5	B
2AR	0	50	50	B
2AS	0	63	37	C
2AT	0	39	61	C
2AU	30	64	6	C
2AV	0	40	60	C
2AW	21	53	26	C
2AX	0	37	63	C
2AY	0	42	58	C
2AZ	0	20	80	C
2BA	0	3	97	C
2BB	0	14	86	C
2BC	0	2	98	C



4. 4th Place

Test Set

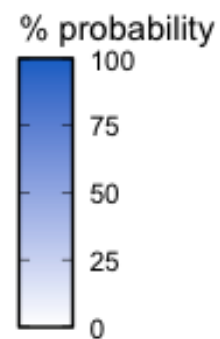
Confusion Matrix



Test Set

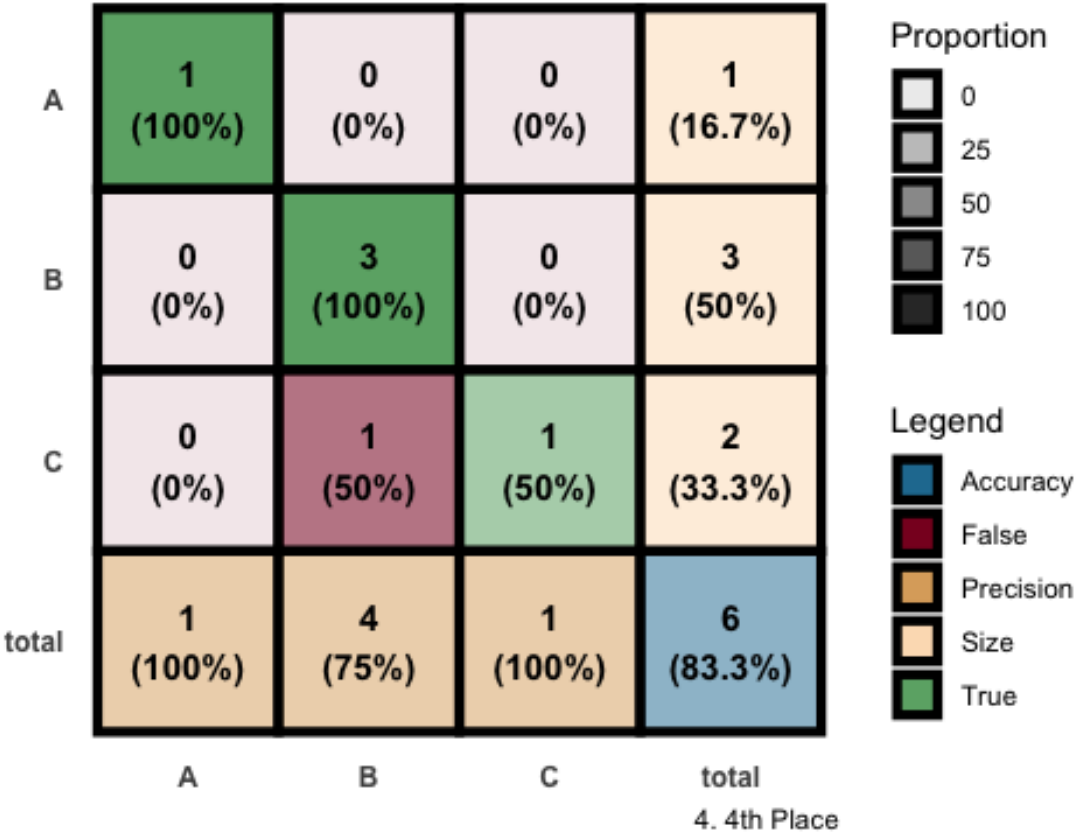
Probability Heatmap

	A	B	C	Exp
2B	100	0	0	A
2C	100	0	0	A
2D	100	0	0	A
2E	100	0	0	A
2F	100	0	0	A
2I	100	0	0	A
2K	100	0	0	A
2N	100	0	0	A
2Q	100	0	0	A
2AD	95	4	1	B
2AI	0	81	19	B
2AL	87	12	1	B
2AN	0	85	15	B



4. 4th Place

External Validation
Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	0	92	8
2BE	0	64	36
2BF	100	0	0
2BG	0	83	17
2BH	0	5	95
2BI	0	51	49

Exp

B

B

A

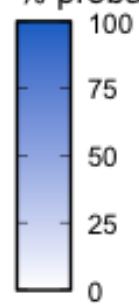
B

C

C

4. 4th Place

% probability



5th Place

Table:

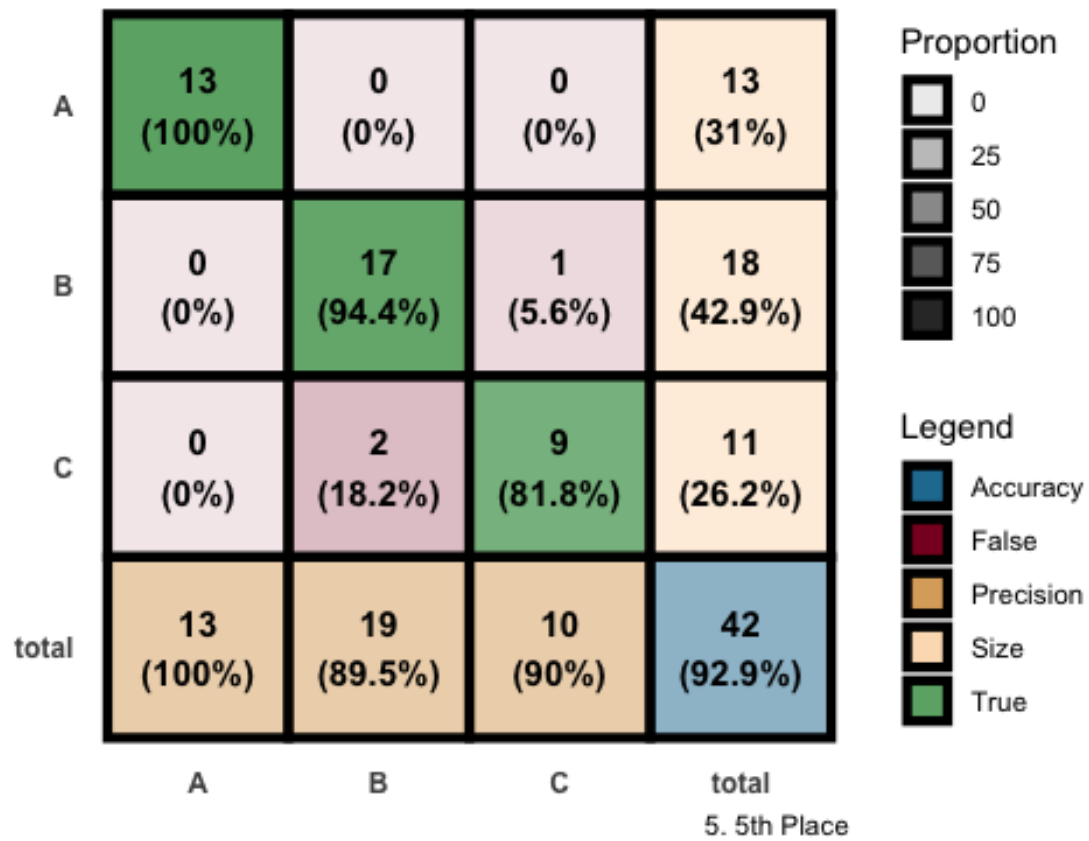
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
92.86%	0.699

Model Coefficients

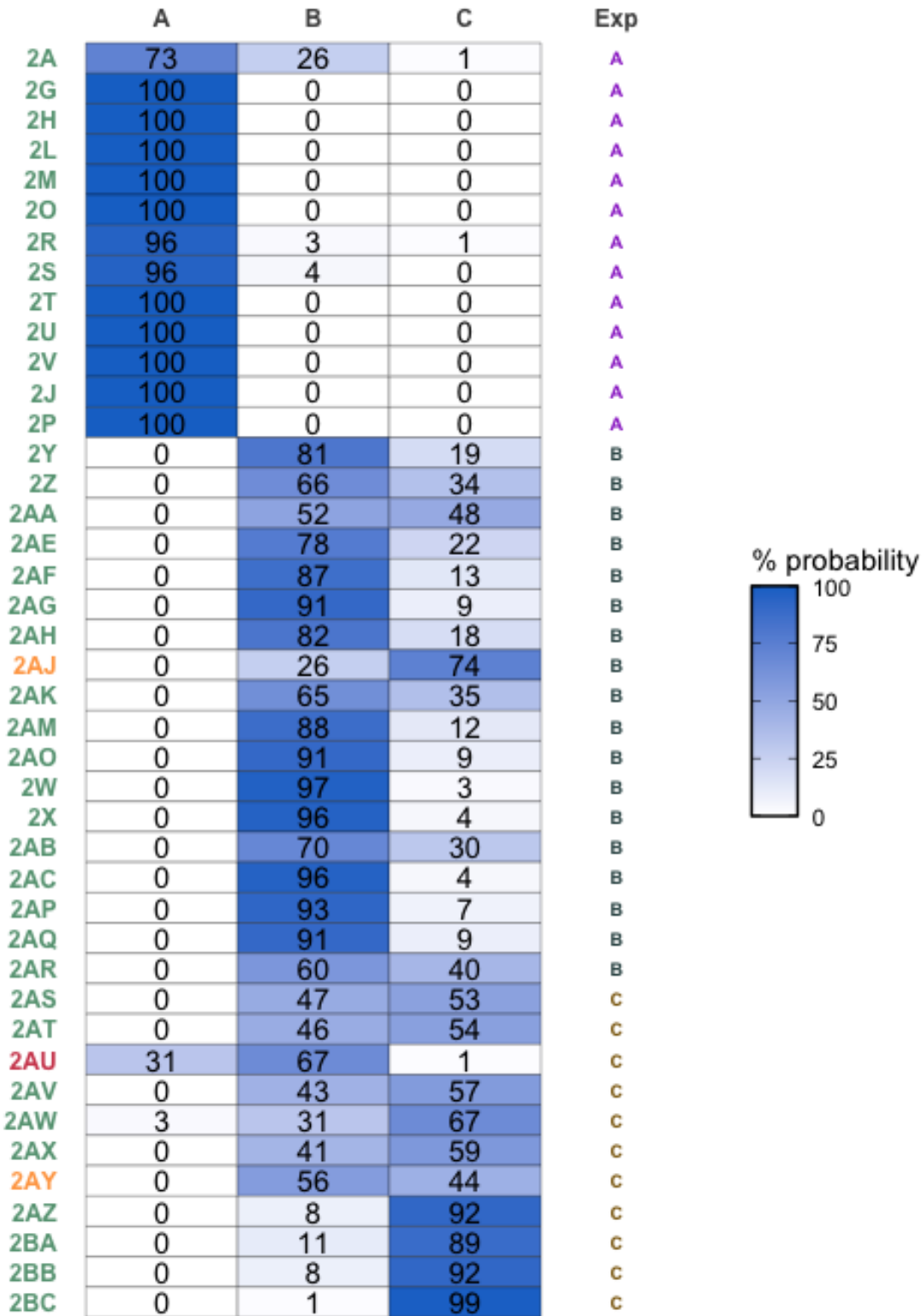
	(Intercept)	dip_x	dip_y	NPA_1	NPA_3
2	8.19396	-23.17355	-36.01353	461.8364	53.46012
3	-120.57703	-23.06822	-36.08947	481.2963	-201.39613

Training Set Confusion Matrix



Training Set

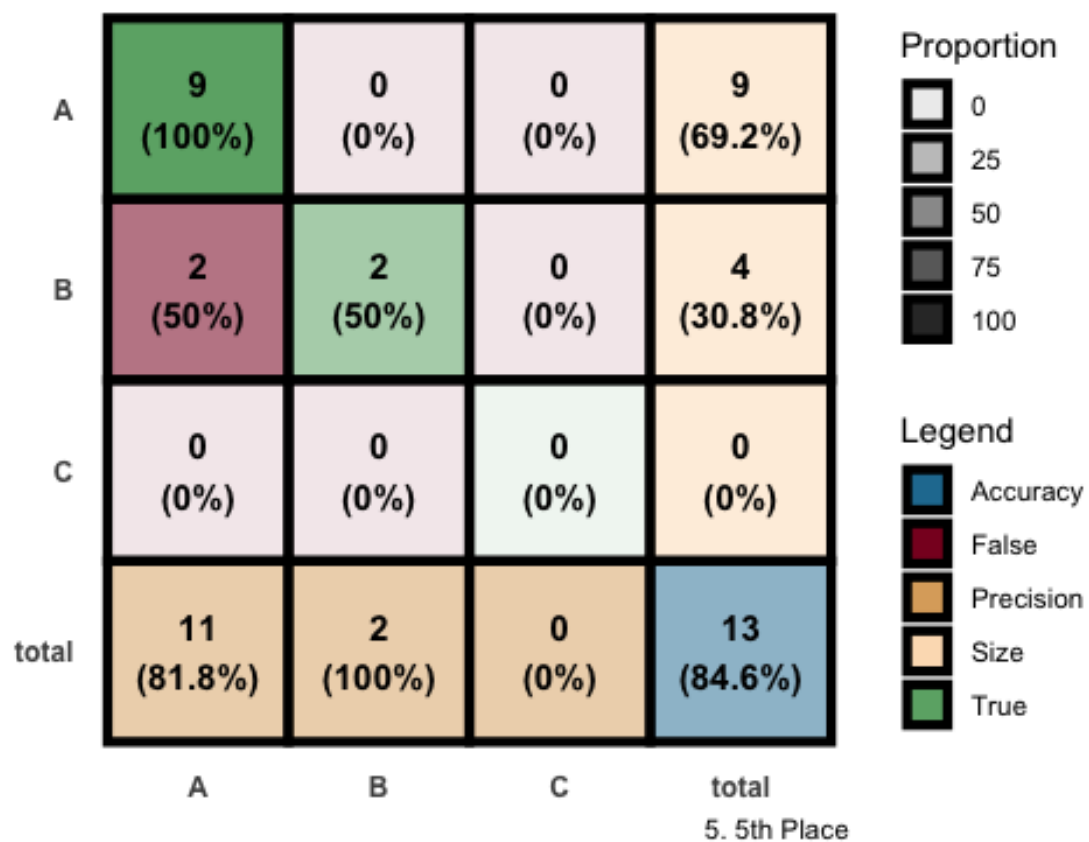
Probability Heatmap



5. 5th Place

Test Set

Confusion Matrix



Test Set

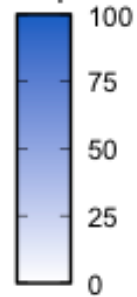
Probability Heatmap

	A	B	C
2B	100	0	0
2C	100	0	0
2D	100	0	0
2E	100	0	0
2F	100	0	0
2I	100	0	0
2K	100	0	0
2N	100	0	0
2Q	100	0	0
2AD	99	1	0
2AI	0	64	36
2AL	52	44	4
2AN	0	89	11

Exp

A
A
A
A
A
A
A
A
A
A
B
B
B
B
B

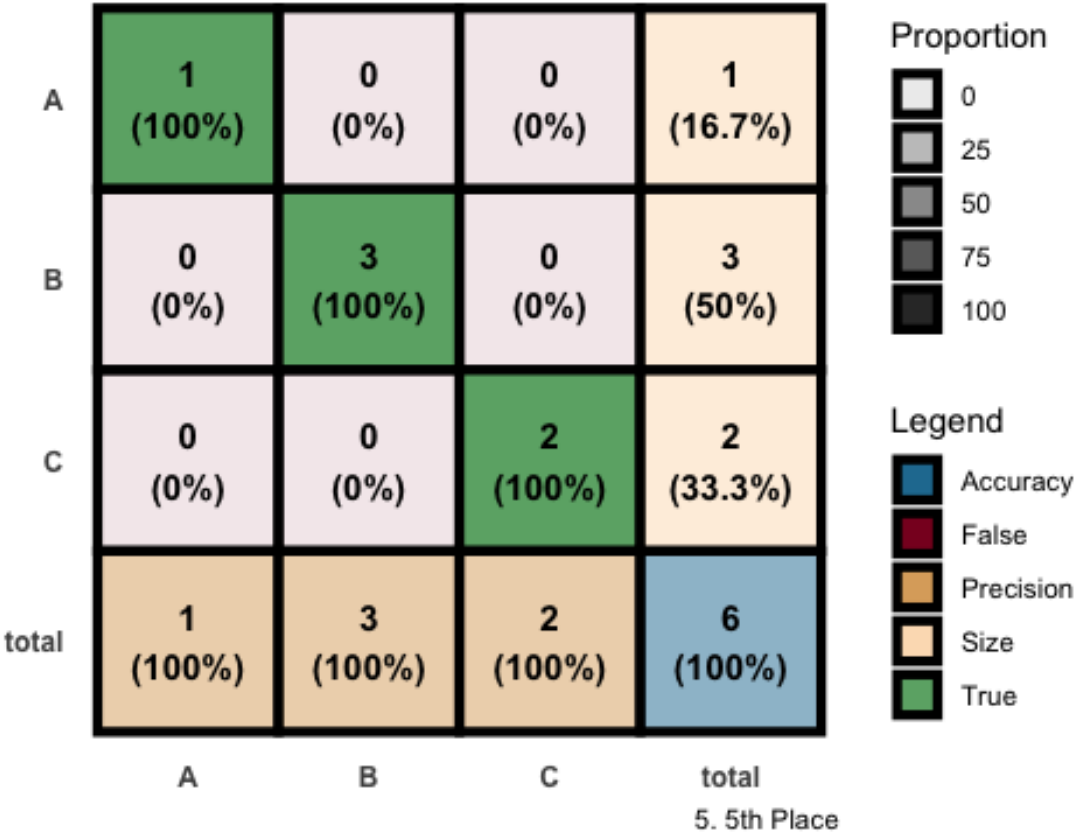
% probability



5. 5th Place

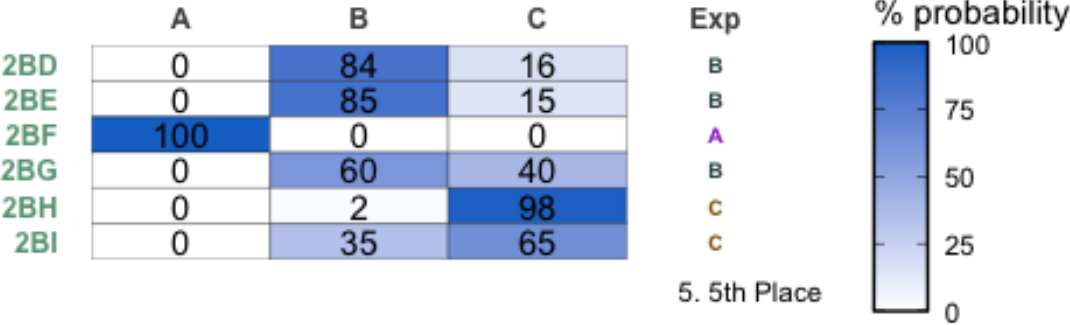
External Validation

Confusion Matrix



External Validation

Probability Heatmap



6th Place

Table:

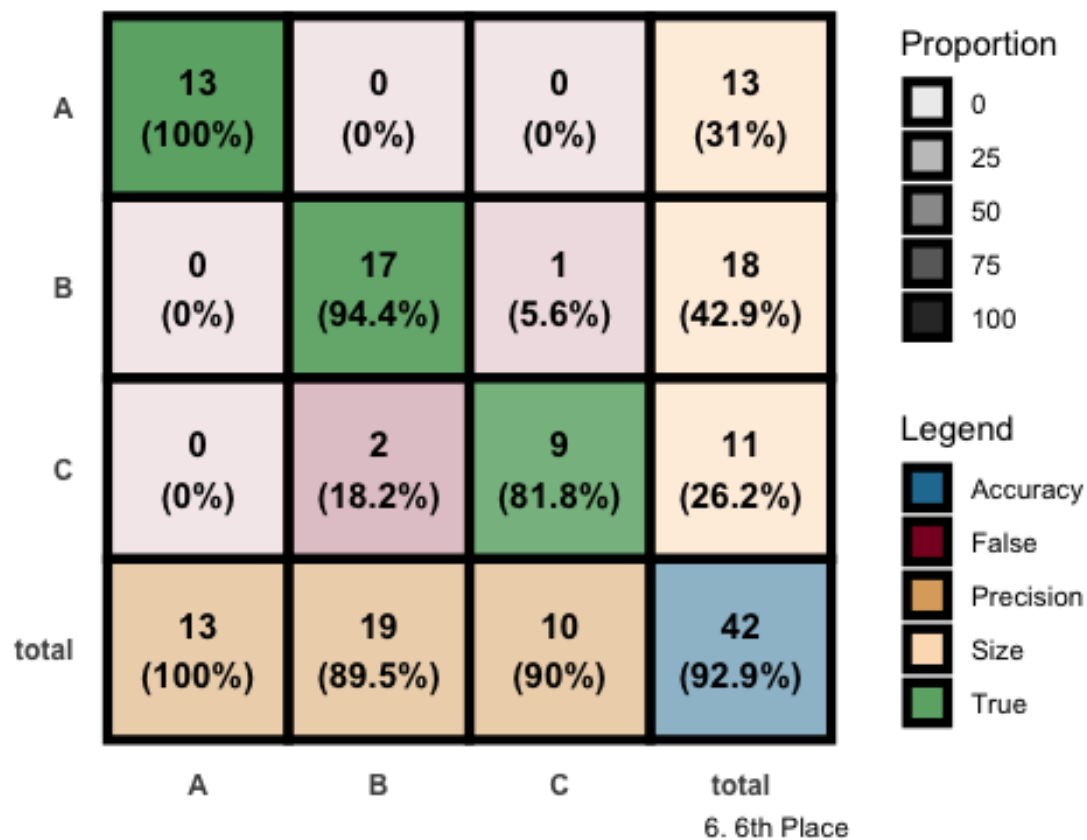
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
92.86%	0.69

Model Coefficients

	(Intercept)	dip_x	dip_y	NPA_3	diff 1 2
2	98.88628	-15.09258	-24.06615	-37.06308	-309.6037
3	-18.63993	-14.97356	-24.24109	-285.15931	-328.8158

Training Set Confusion Matrix

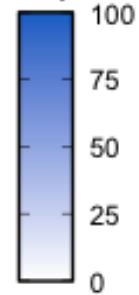


Training Set

Probability Heatmap

	A	B	C	Exp
2A	56	43	1	A
2G	100	0	0	A
2H	100	0	0	A
2L	100	0	0	A
2M	100	0	0	A
2O	100	0	0	A
2R	94	5	1	A
2S	94	6	0	A
2T	100	0	0	A
2U	100	0	0	A
2V	100	0	0	A
2J	100	0	0	A
2P	100	0	0	A
2Y	0	79	21	B
2Z	0	64	36	B
2AA	0	51	49	B
2AE	0	79	21	B
2AF	0	89	11	B
2AG	0	91	9	B
2AH	0	83	17	B
2AJ	0	27	73	B
2AK	0	66	34	B
2AM	0	89	11	B
2AO	0	91	9	B
2W	2	95	3	B
2X	0	95	5	B
2AB	0	71	29	B
2AC	0	95	5	B
2AP	0	93	7	B
2AQ	0	91	9	B
2AR	0	60	40	B
2AS	0	46	54	C
2AT	0	45	55	C
2AU	46	53	1	C
2AV	0	43	57	C
2AW	8	31	61	C
2AX	0	41	59	C
2AY	0	56	44	C
2AZ	0	8	92	C
2BA	0	6	94	C
2BB	0	5	95	C
2BC	0	1	99	C

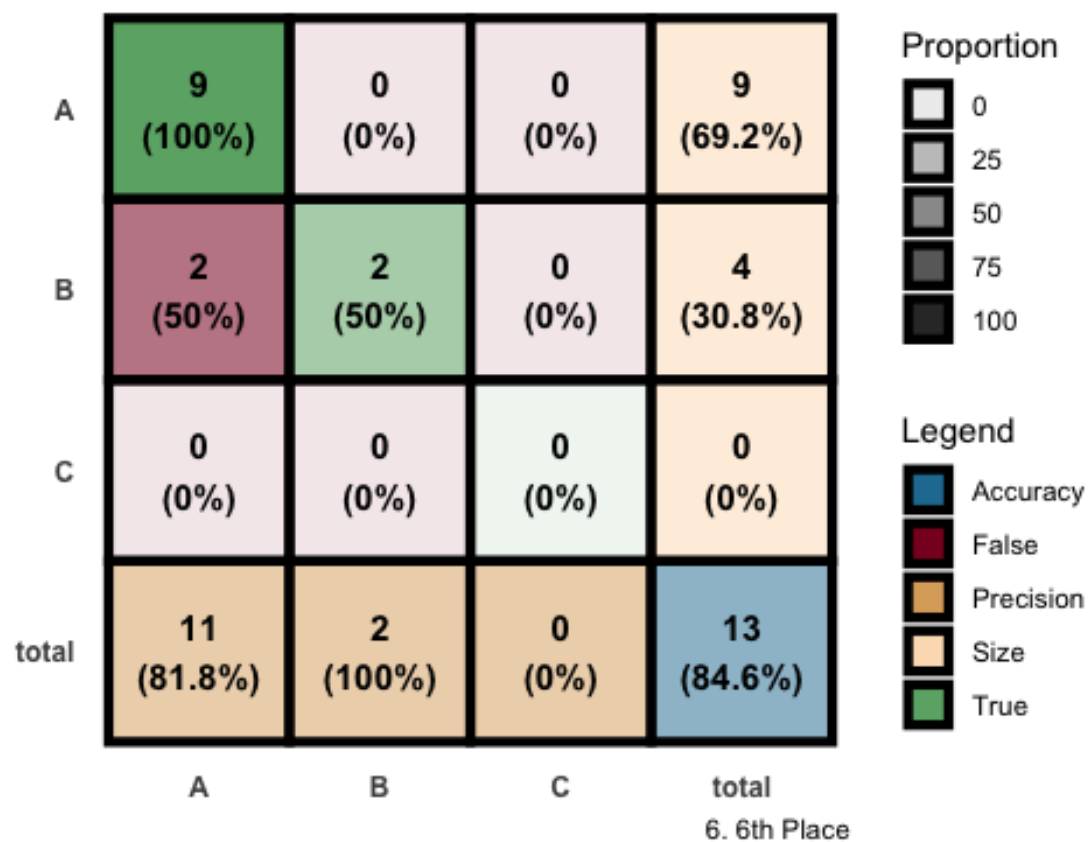
% probability



6. 6th Place

Test Set

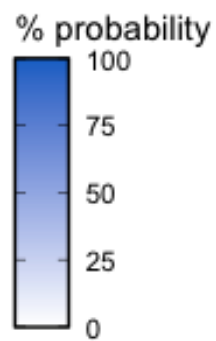
Confusion Matrix



Test Set

Probability Heatmap

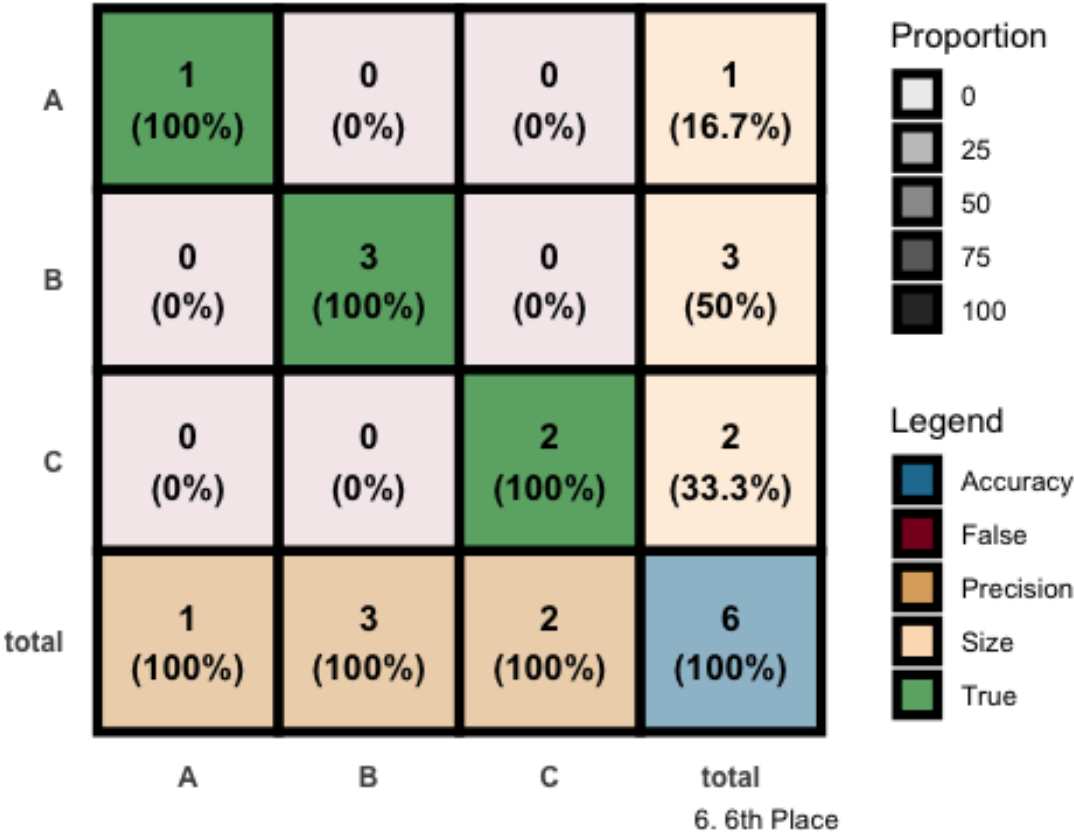
	A	B	C	Exp
2B	100	0	0	A
2C	100	0	0	A
2D	100	0	0	A
2E	100	0	0	A
2F	100	0	0	A
2I	100	0	0	A
2K	100	0	0	A
2N	100	0	0	A
2Q	100	0	0	A
2AD	90	10	0	B
2AI	0	66	34	B
2AL	85	14	1	B
2AN	0	88	12	B



6. 6th Place

External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	0	86	14
2BE	0	83	17
2BF	100	0	0
2BG	0	62	38
2BH	0	2	98
2BI	0	35	65

Exp

B

B

A

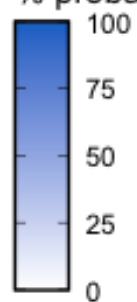
B

C

C

6. 6th Place

% probability



7th Place

Table:

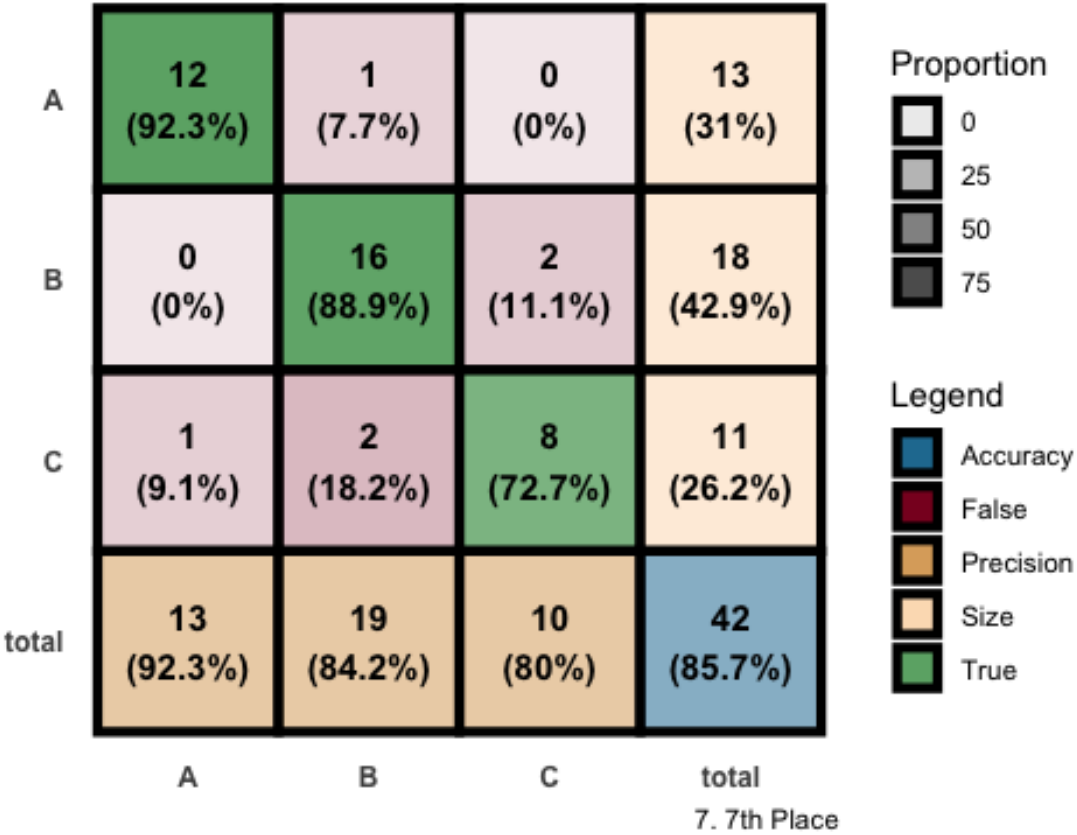
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.687

Model Coefficients

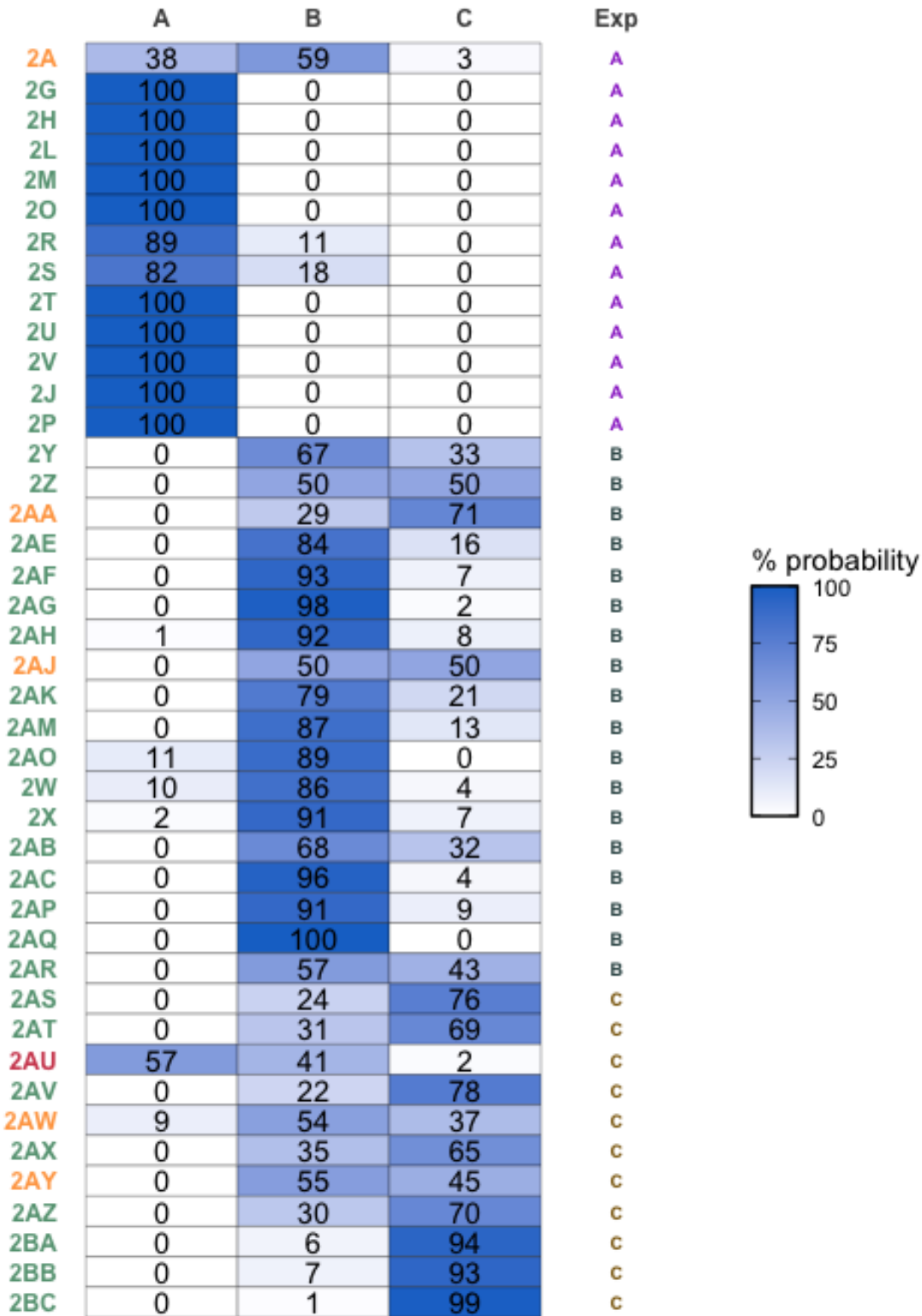
	(Intercept)	Total	NPA_1	NPA_3	loc.B5
2	-227.6700	13.29595	134.6280	-434.2698	-2.661891
3	-339.3368	13.62434	149.4287	-654.3205	-3.123991

Training Set
Confusion Matrix



Training Set

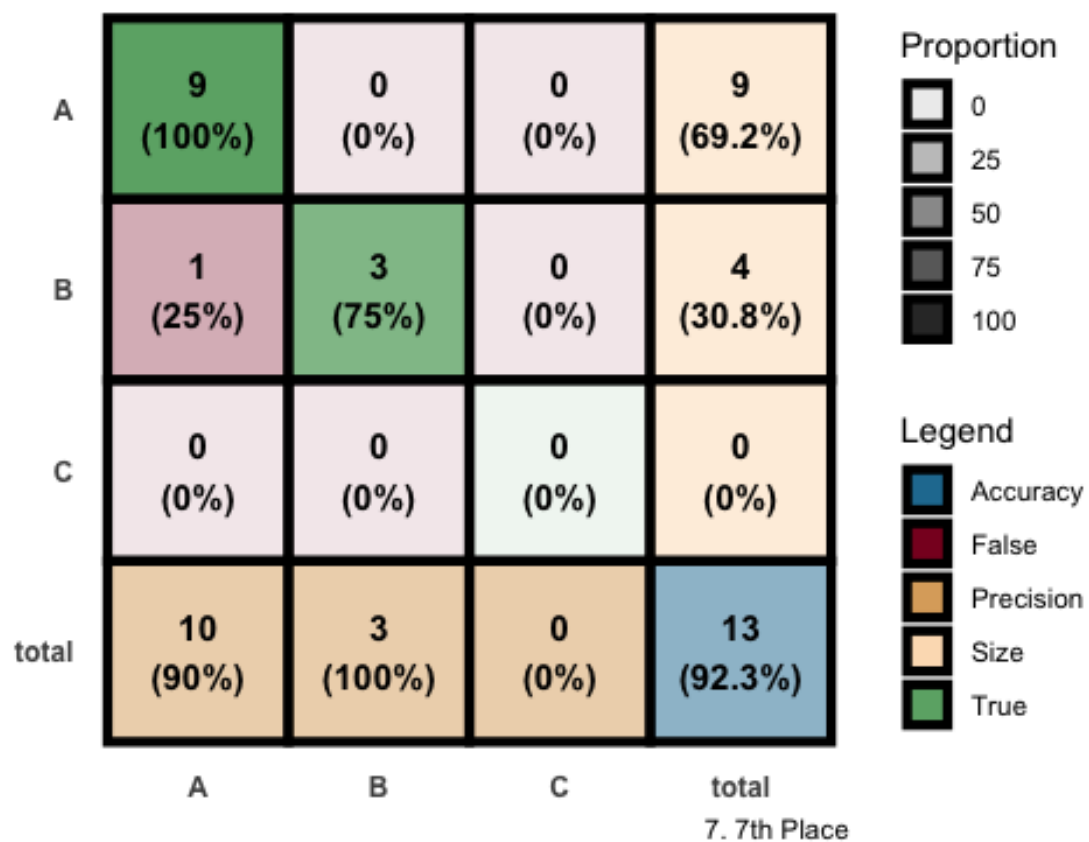
Probability Heatmap



7. 7th Place

Test Set

Confusion Matrix



Test Set

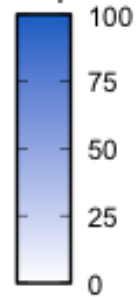
Probability Heatmap

	A	B	C
2B	100	0	0
2C	100	0	0
2D	100	0	0
2E	100	0	0
2F	100	0	0
2I	100	0	0
2K	100	0	0
2N	100	0	0
2Q	100	0	0
2AD	6	87	7
2AI	0	89	11
2AL	100	0	0
2AN	0	97	3

Exp

A
A
A
A
A
A
A
A
A
A
B
B
B
B
B

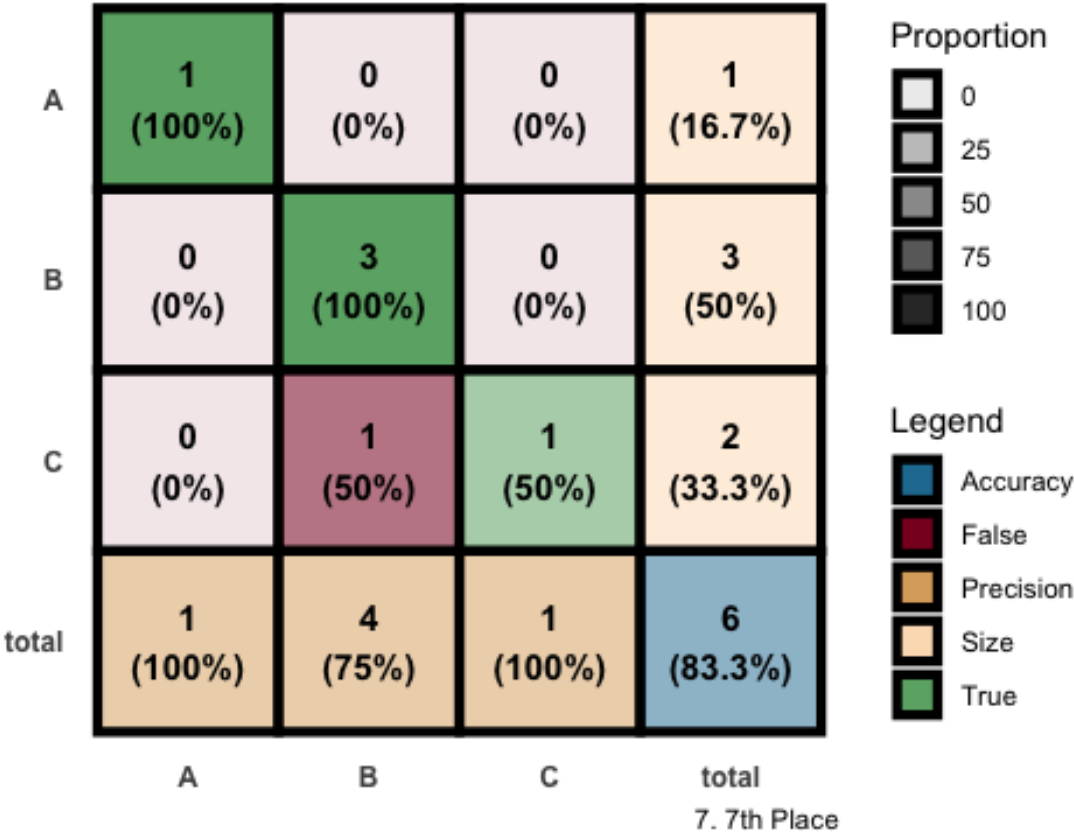
% probability



7. 7th Place

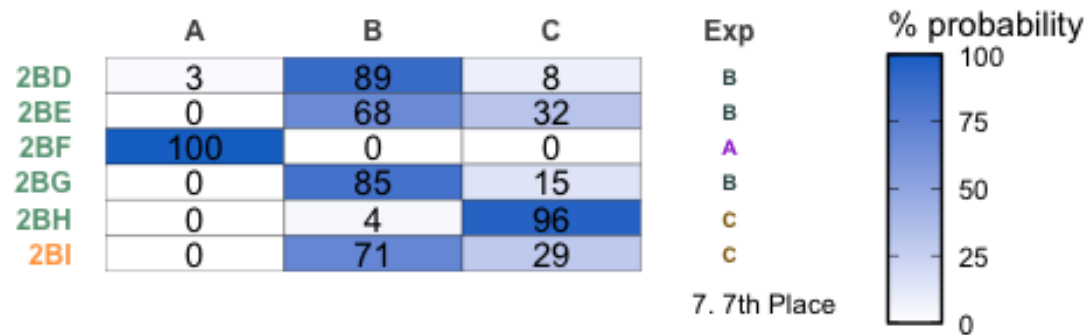
External Validation

Confusion Matrix



External Validation

Probability Heatmap



8th Place

Table:

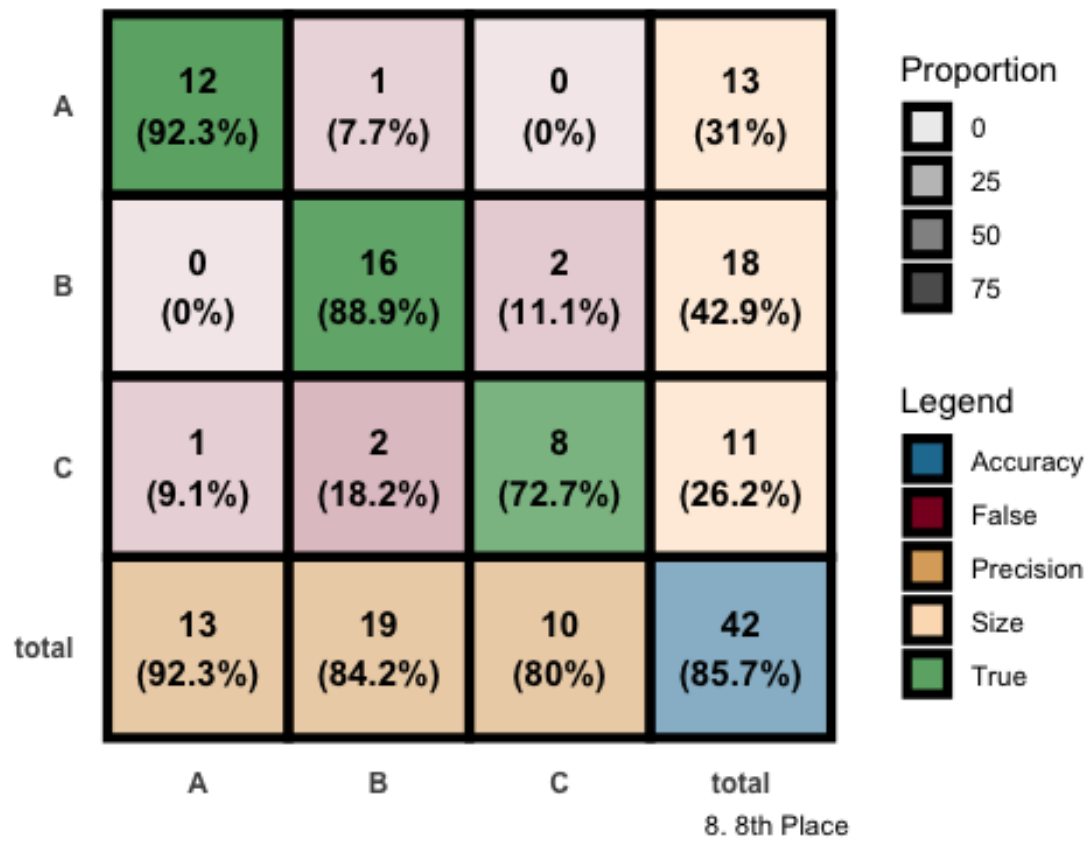
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.678

Model Coefficients

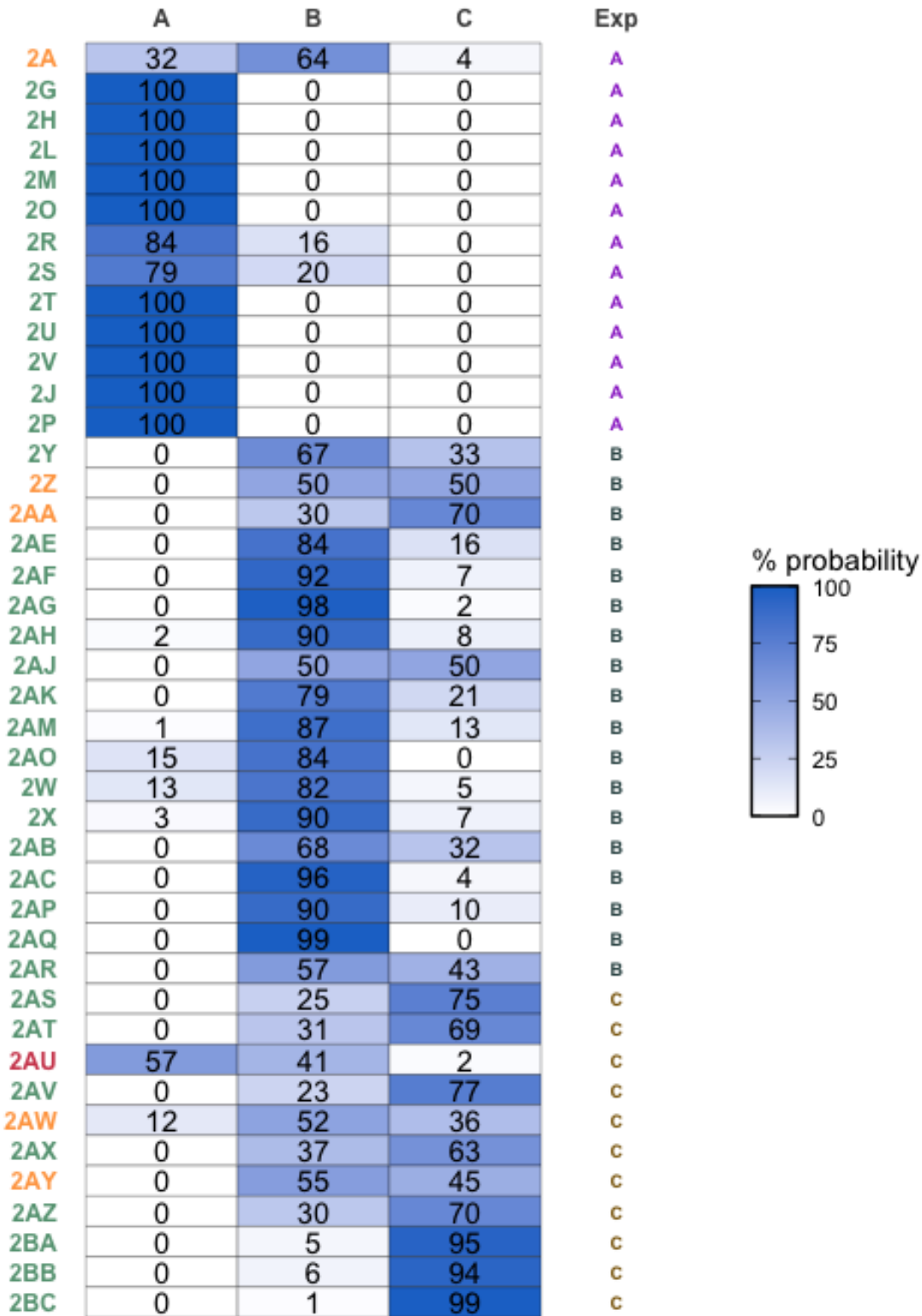
	(Intercept)	Total	NPA_3	diff 1 2	loc.B5
2	-152.6517	10.99773	-377.3616	-108.4889	-2.136966
3	-255.7061	11.31747	-591.6304	-122.4150	-2.558339

Training Set Confusion Matrix



Training Set

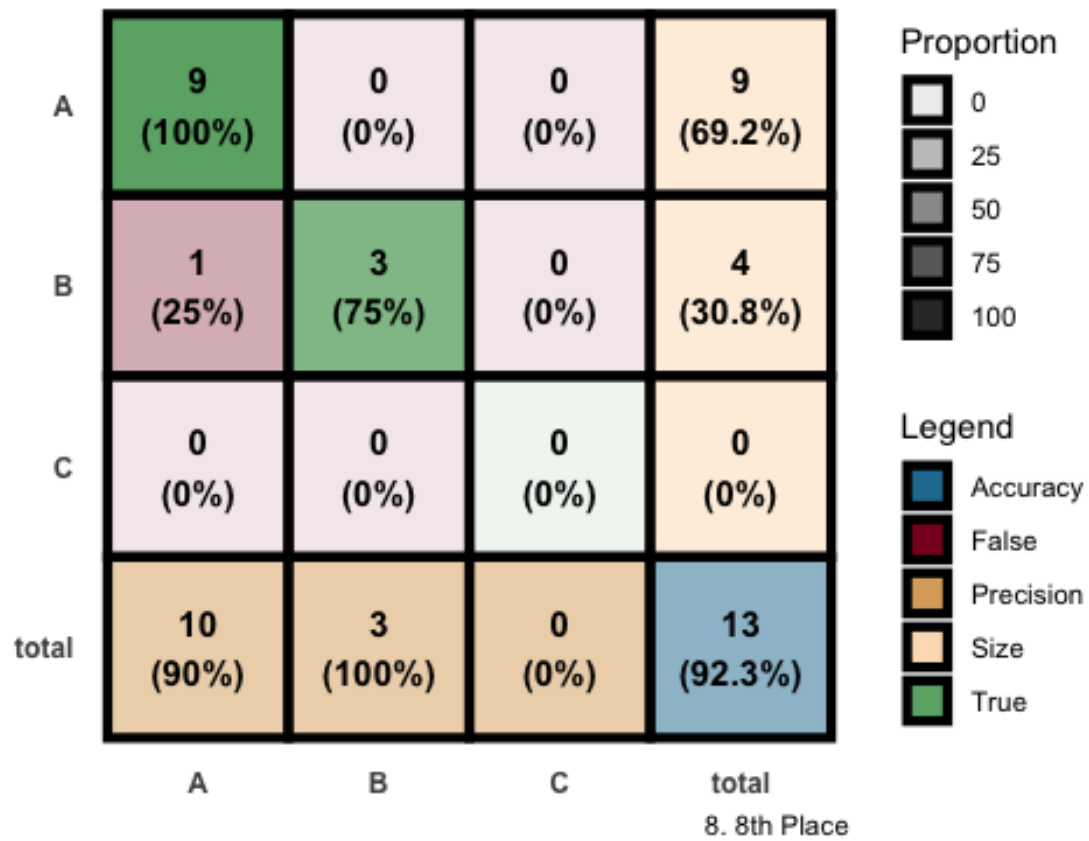
Probability Heatmap



8. 8th Place

Test Set

Confusion Matrix



Test Set

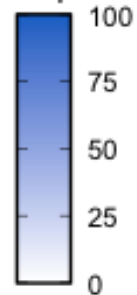
Probability Heatmap

	A	B	C
2B	100	0	0
2C	100	0	0
2D	100	0	0
2E	100	0	0
2F	100	0	0
2I	100	0	0
2K	100	0	0
2N	100	0	0
2Q	100	0	0
2AD	8	85	7
2AI	0	88	11
2AL	100	0	0
2AN	0	96	3

Exp

A
A
A
A
A
A
A
A
A
A
B
B
B
B
B

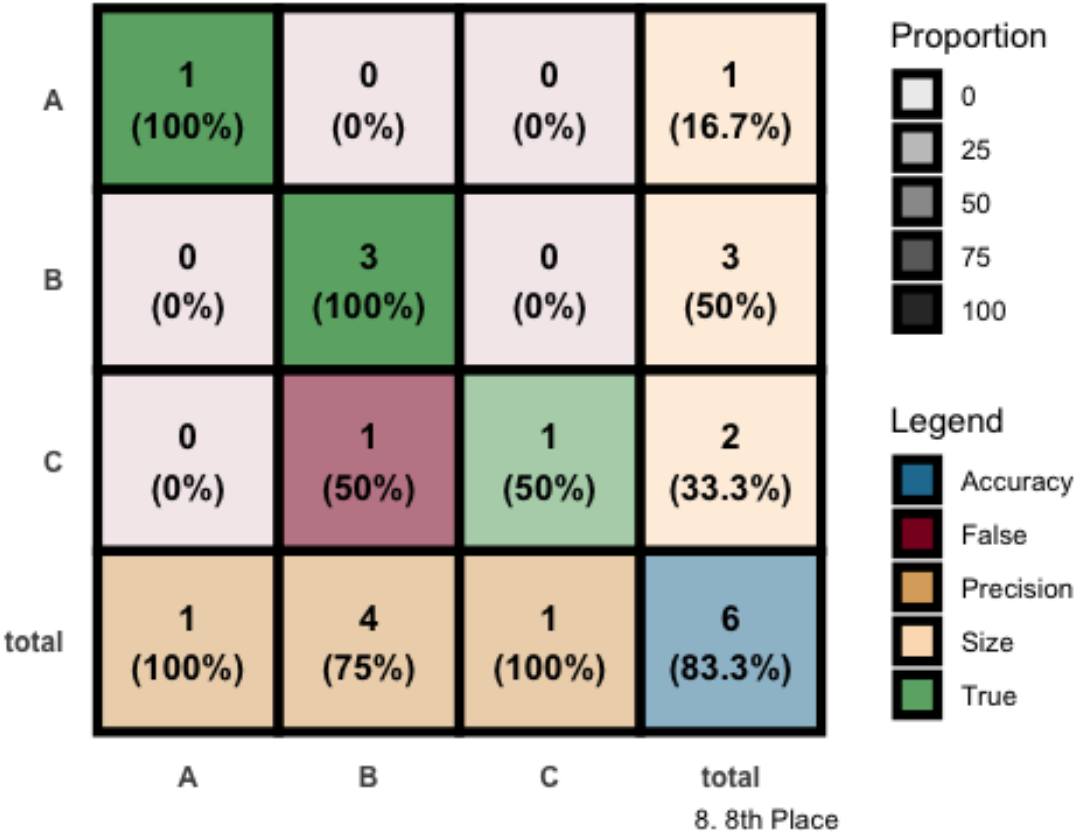
% probability



8. 8th Place

External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	8	84	8
2BE	0	67	33
2BF	100	0	0
2BG	0	83	17
2BH	0	4	96
2BI	0	71	29

Exp

B

B

A

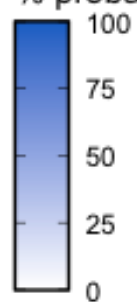
B

C

C

8. 8th Place

% probability



9th Place

Table:

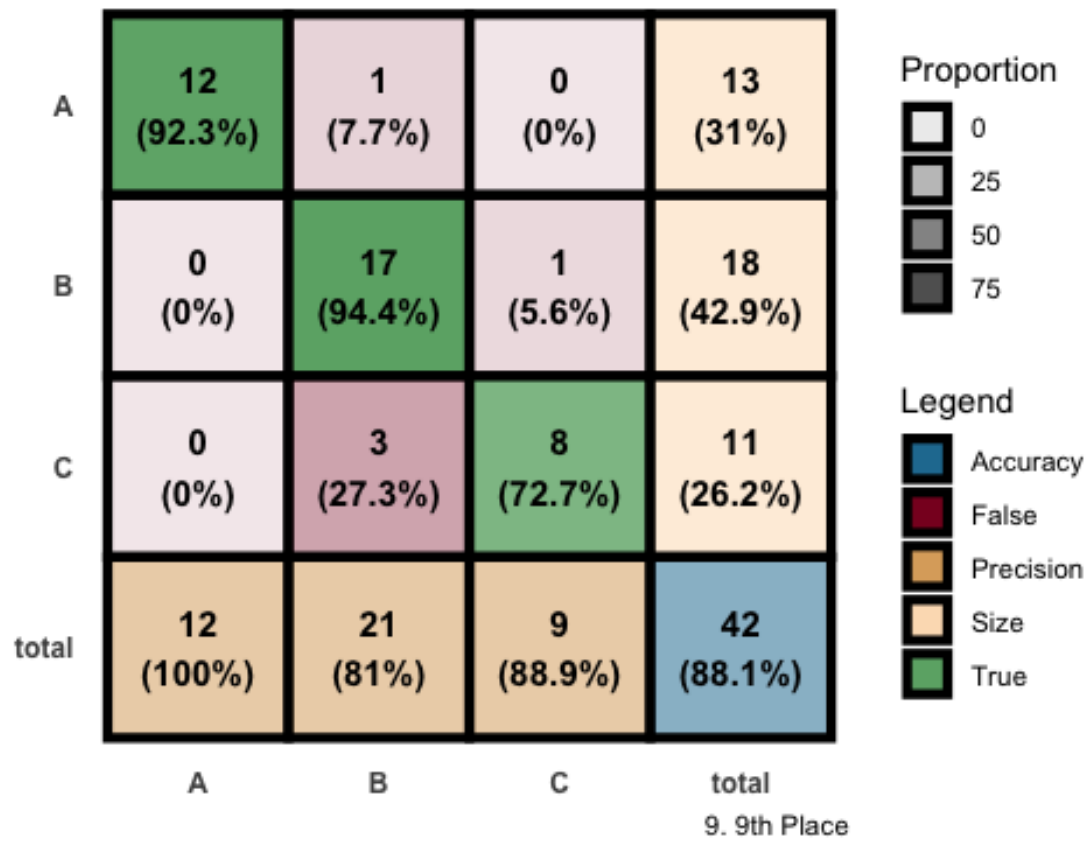
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
88.1%	0.672

Model Coefficients

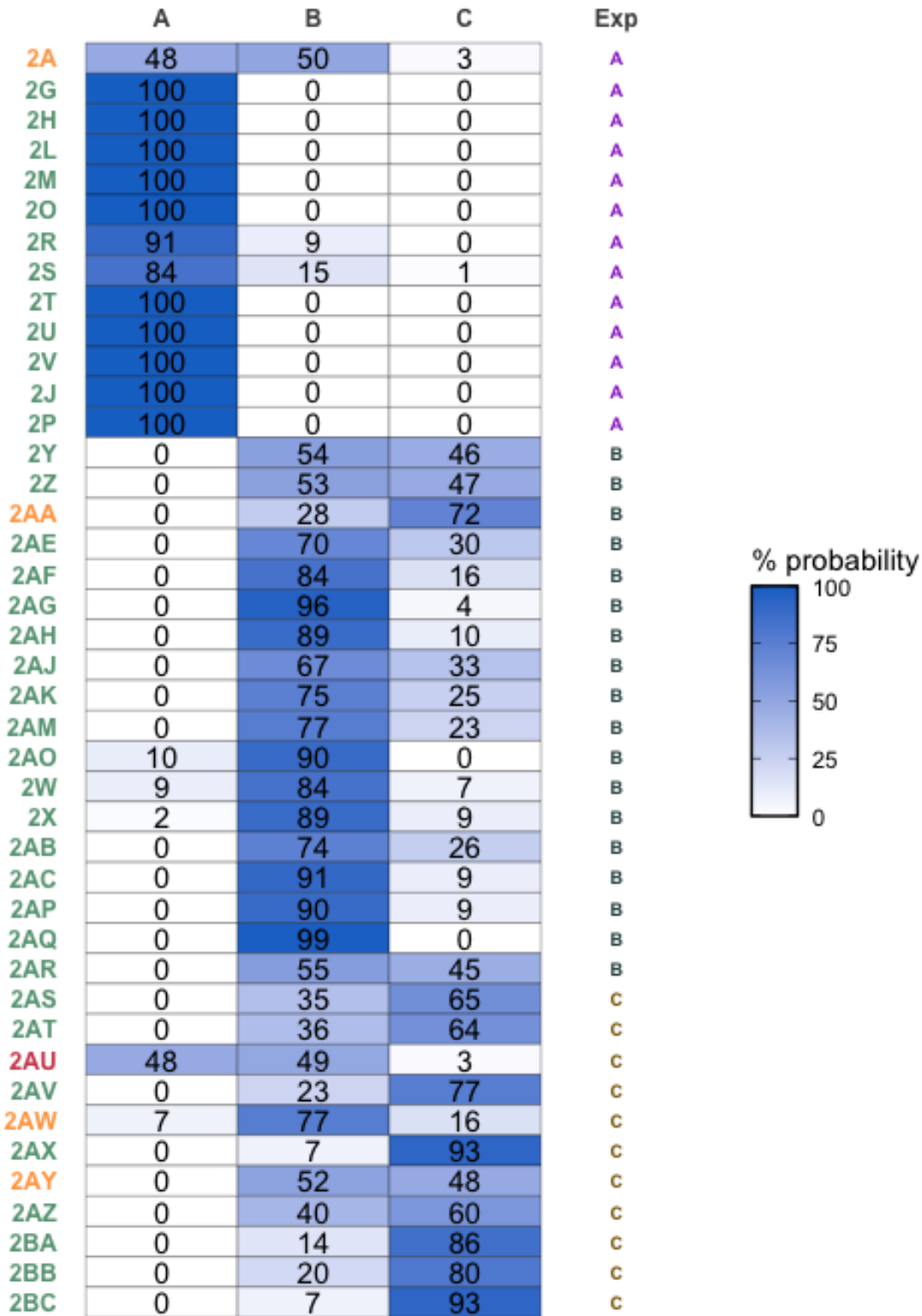
	(Intercept)	Total	diff 1 2	diff 2 3	loc.B5
2	-329.4002	13.23035	-136.3753	410.1141	-2.819057
3	-443.3563	14.33282	-159.1235	543.2315	-3.494759

Training Set Confusion Matrix



Training Set

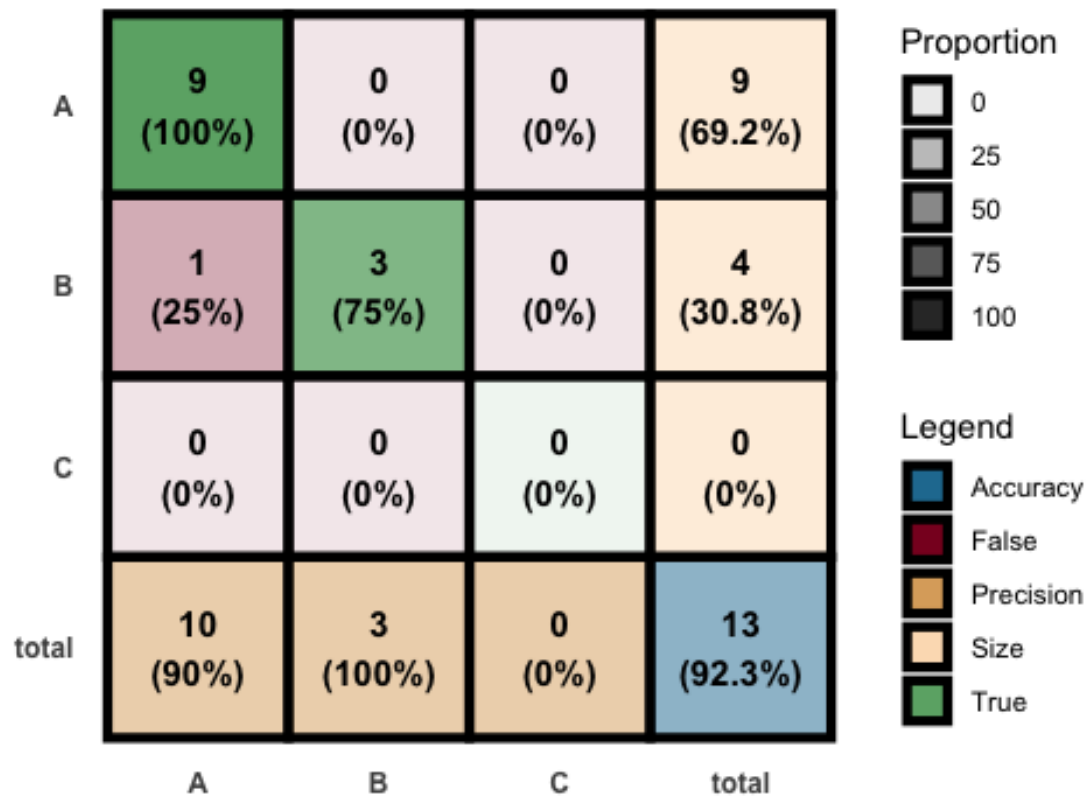
Probability Heatmap



9. 9th Place

Test Set

Confusion Matrix



9. 9th Place

Test Set

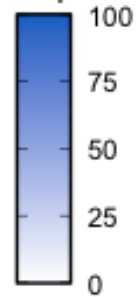
Probability Heatmap

	A	B	C
2B	100	0	0
2C	100	0	0
2D	100	0	0
2E	100	0	0
2F	100	0	0
2I	100	0	0
2K	100	0	0
2N	100	0	0
2Q	100	0	0
2AD	8	84	8
2AI	0	90	10
2AL	100	0	0
2AN	0	92	8

Exp

A
A
A
A
A
A
A
A
A
A
B
B
B
B

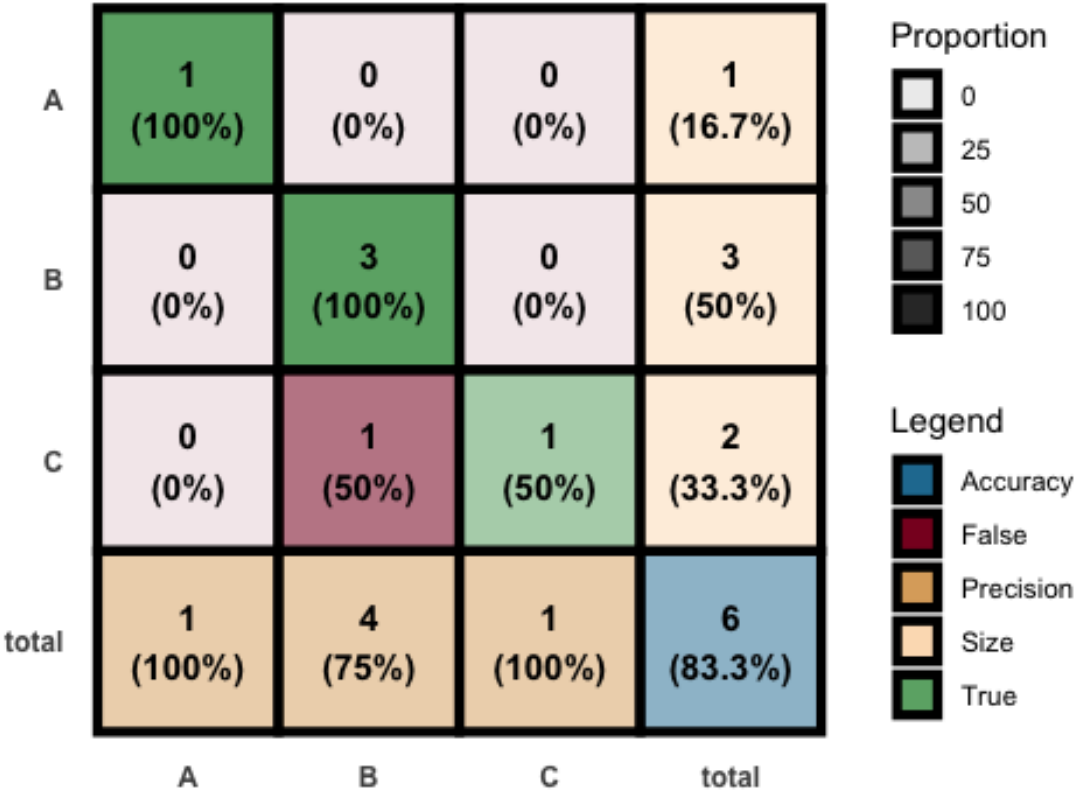
% probability



9. 9th Place

External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	1	88	11
2BE	0	80	20
2BF	100	0	0
2BG	0	87	13
2BH	0	5	95
2BI	0	68	32

Exp

B

B

A

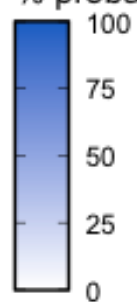
B

C

C

9. 9th Place

% probability



10th Place

Table:

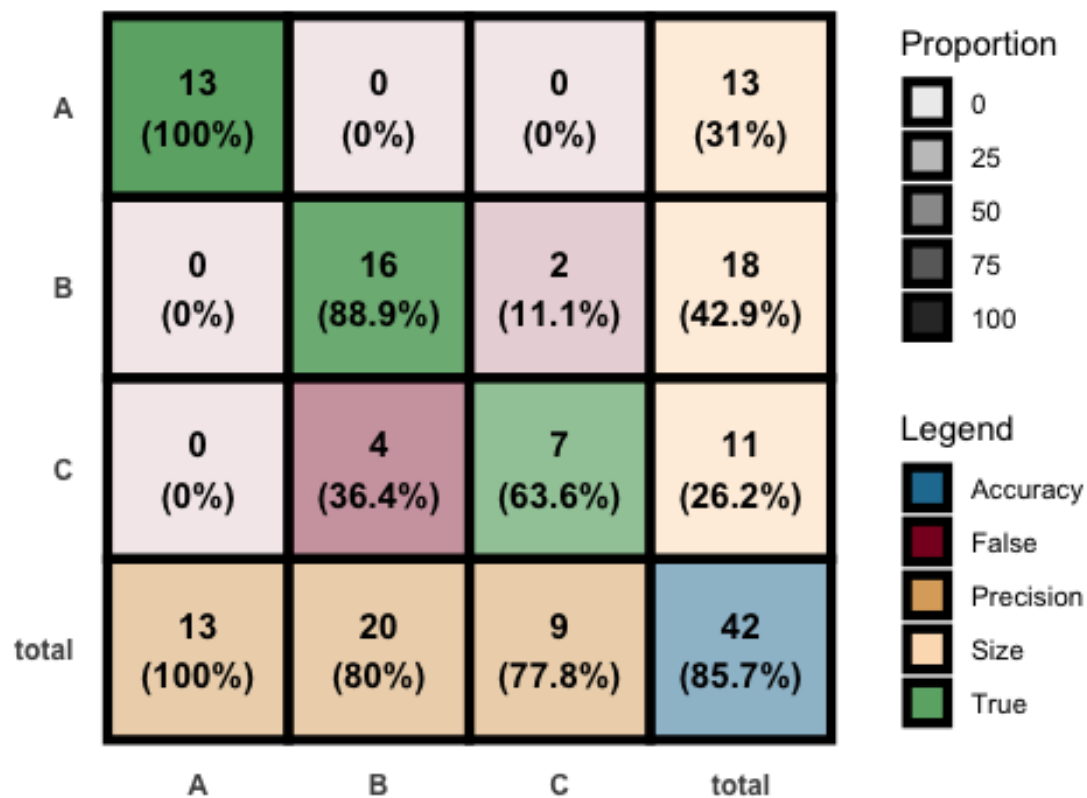
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.671

Model Coefficients

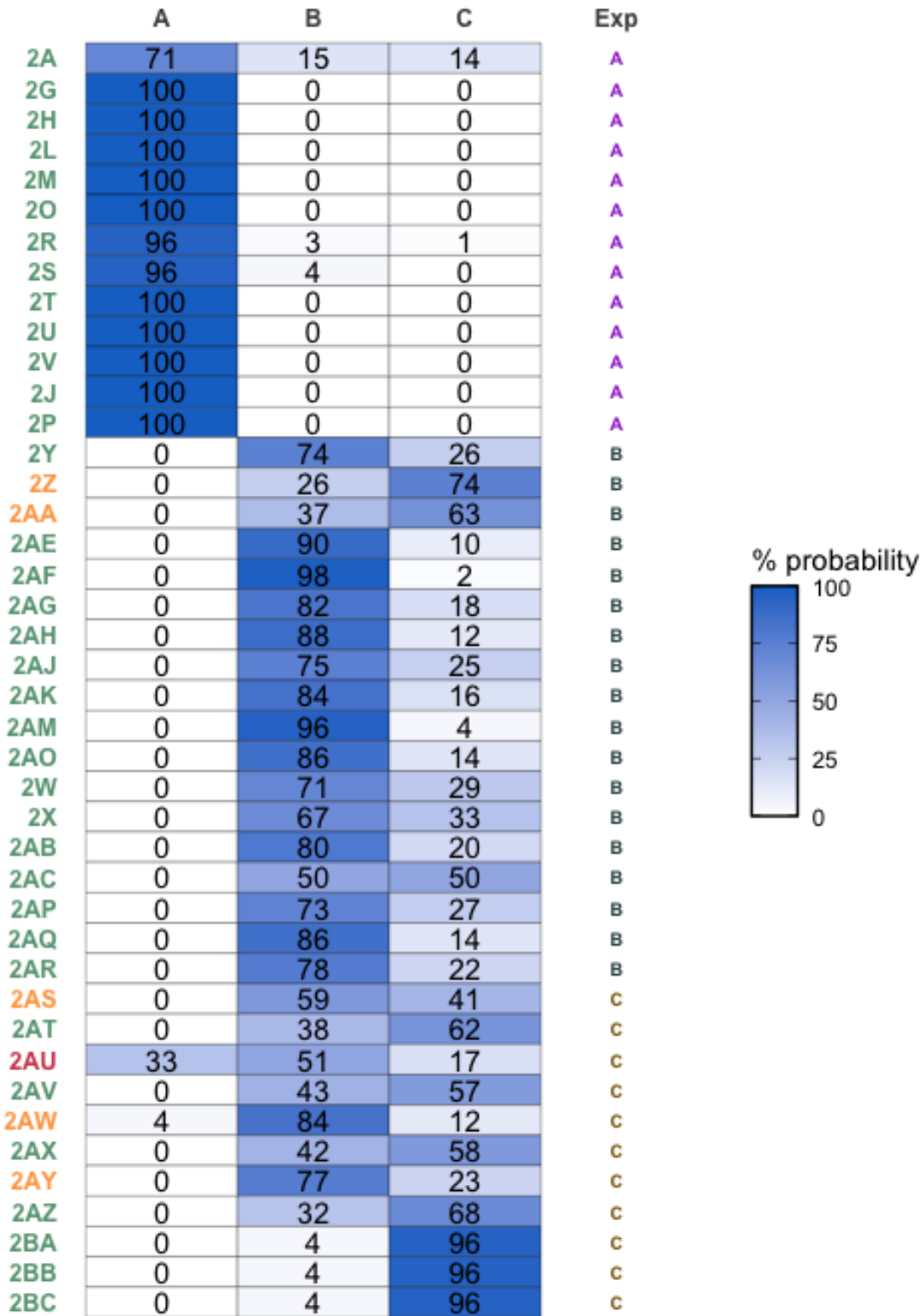
	(Intercept)	dip_x	dip_y	NPA_1	diff 1 2
2	23.03156	-25.57605	-40.08803	428.3814	-98.56926
3	105.48531	-25.79453	-41.40318	208.4211	-311.07207

Training Set Confusion Matrix



Training Set

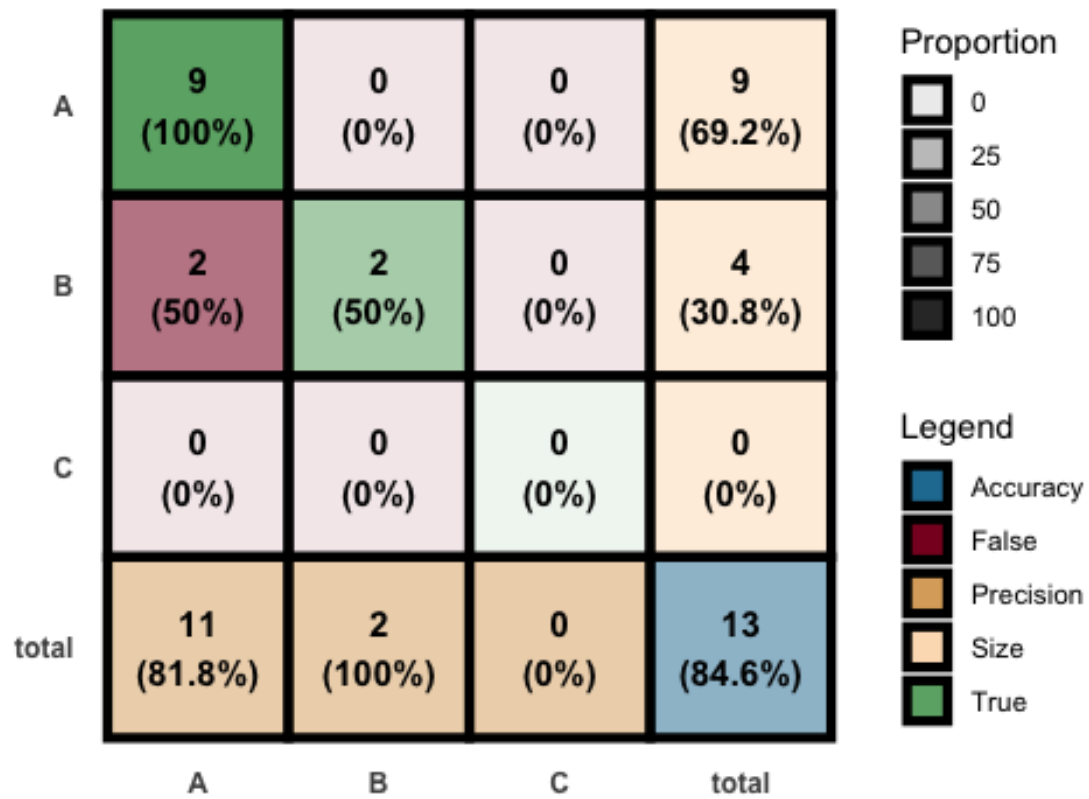
Probability Heatmap



10. 10th Place

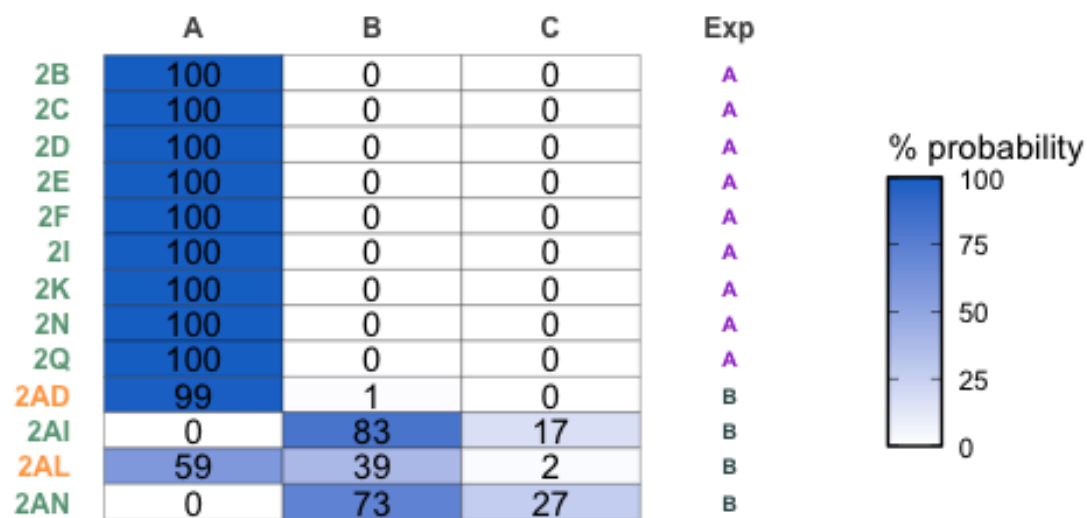
Test Set

Confusion Matrix



Test Set

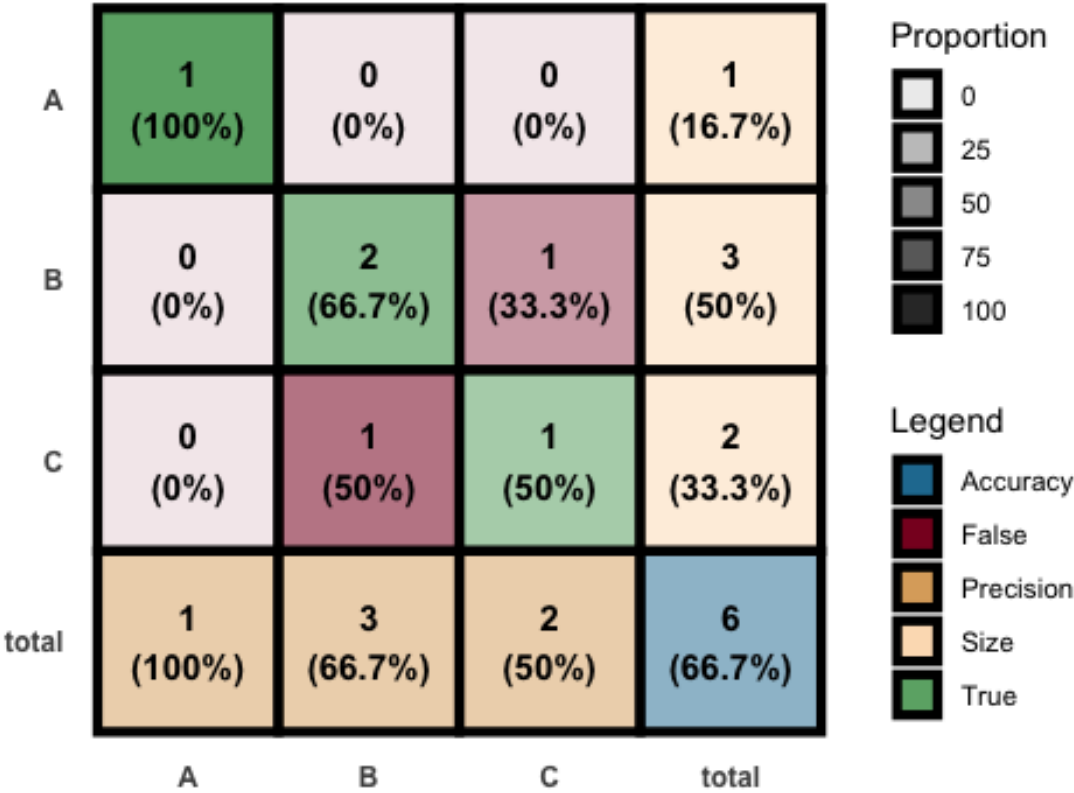
Probability Heatmap



10. 10th Place

External Validation

Confusion Matrix



External Validation

Probability Heatmap

