

Aldehyde Deuteration - Model Search

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_y + Dist(2, 7) + NPA_1	0.577
class ~ -2-3- + dip_y + NPA_1 + NPA_3	0.577
class ~ dip_y + NPA_3 + diff 1 2 + L	0.570
class ~ dip_y + Dist(2, 7) + NPA_3 + diff 1 2	0.567
class ~ -2-3- + dip_y + NPA_1 + NPA_2	0.557
class ~ -2-3- + dip_y + Dist(2, 7) + diff 1 2	0.555
class ~ -2-7- + dip_y + NPA_1 + NPA_3	0.555
class ~ dip_y + NPA_1 + NPA_3 + diff 2 7	0.553
class ~ dip_y + NPA_3 + NPA_7 + diff 1 2	0.551
class ~ -2-3- + Total + NPA_1 + NPA_3	0.549

Section Results Summary

	Training Accuracy	Test Accuracy
1. 1st Place	90.48	92.3
2. 2nd Place	85.71	100.0
3. 3rd Place	78.57	92.3
4. 4th Place	85.71	100.0
5. 5th Place	80.95	92.3
6. 6th Place	85.71	92.3
7. 7th Place	83.33	100.0
8. 8th Place	83.33	100.0

	Training Accuracy	Test Accuracy
9. 9th Place	83.33	100.0
10. 10th Place	85.71	84.6

1st Place

Table:

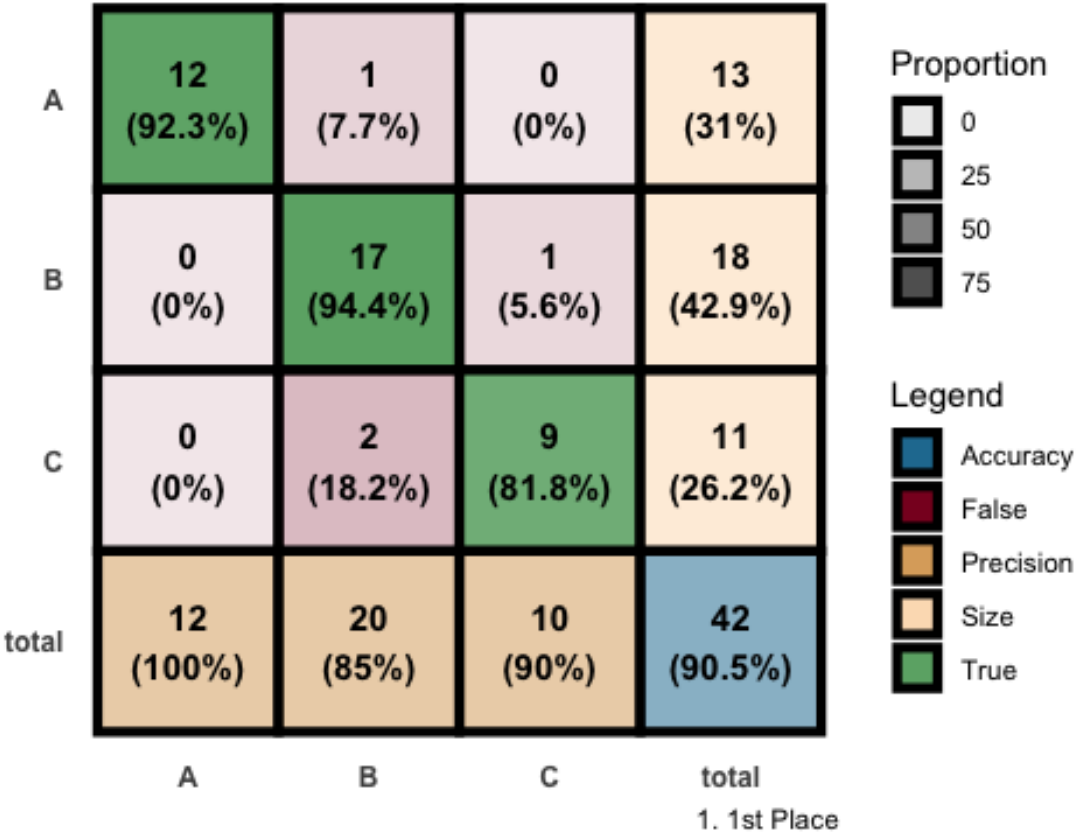
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
90.48%	0.577

Model Coefficients

	x
-2-3-	-0.3801322
dip_y	-1.5680528
Dist(2, 7)	295.8494658
NPA_1	47.5647077

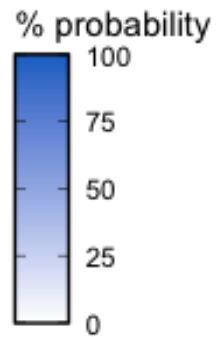
Training Set
Confusion Matrix



Training Set

Probability Heatmap

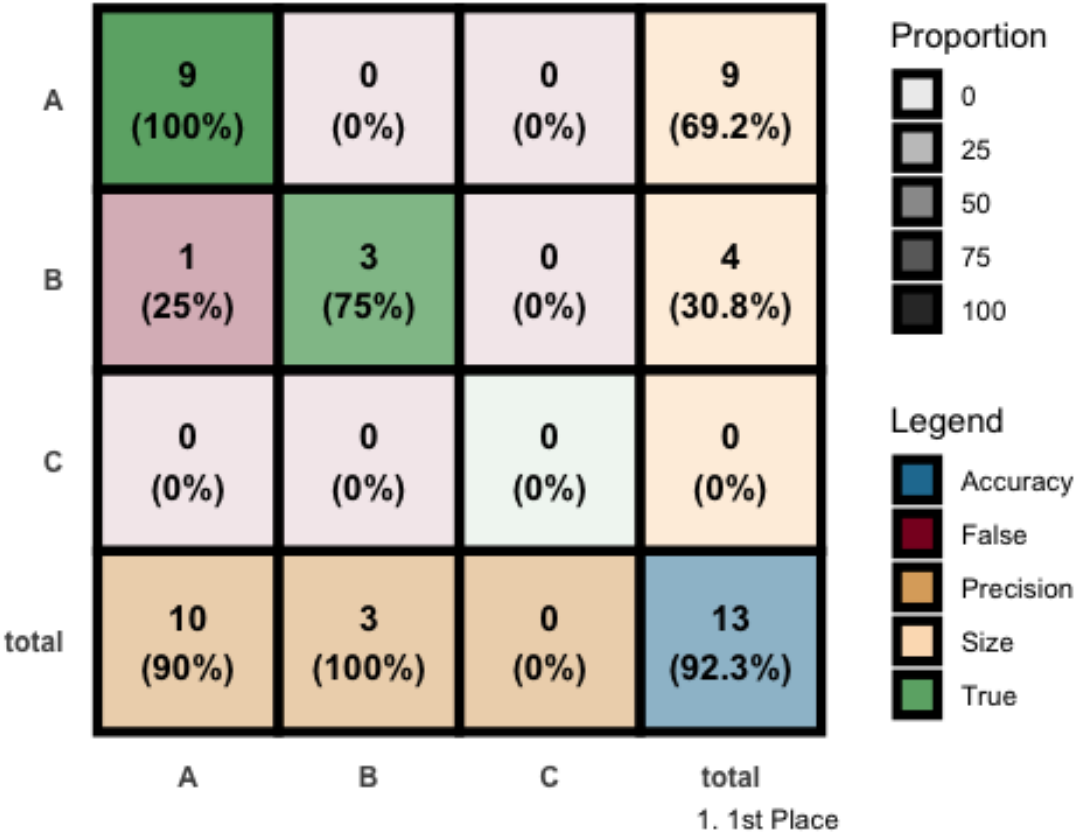
	A	B	C	Exp
2A	57	42	1	A
2G	98	2	0	A
2H	98	2	0	A
2L	88	11	0	A
2M	89	11	0	A
2O	100	0	0	A
2R	13	80	6	A
2S	71	28	0	A
2T	74	25	0	A
2U	99	1	0	A
2V	82	18	0	A
2J	87	13	0	A
2P	98	2	0	A
2Y	2	68	30	B
2Z	2	65	33	B
2AA	1	51	48	B
2AE	10	81	8	B
2AF	46	53	1	B
2AG	18	77	4	B
2AH	16	79	5	B
2AJ	2	62	36	B
2AK	6	81	13	B
2AM	7	81	13	B
2AO	23	74	3	B
2W	27	70	3	B
2X	16	79	5	B
2AB	3	71	26	B
2AC	5	79	15	B
2AP	20	76	4	B
2AQ	22	75	4	B
2AR	0	24	76	B
2AS	1	44	55	C
2AT	1	44	55	C
2AU	38	60	2	C
2AV	1	46	54	C
2AW	5	79	16	C
2AX	0	2	98	C
2AY	0	23	77	C
2AZ	0	2	98	C
2BA	0	0	100	C
2BB	0	0	100	C
2BC	0	0	100	C



1. 1st Place

Test Set

Confusion Matrix



Test Set

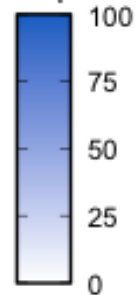
Probability Heatmap

	A	B	C
2B	86	14	0
2C	85	15	0
2D	100	0	0
2E	83	17	0
2F	81	19	0
2I	92	8	0
2K	100	0	0
2N	77	23	0
2Q	98	2	0
2AD	29	69	2
2AI	5	78	17
2AL	67	33	1
2AN	10	81	8

Exp

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

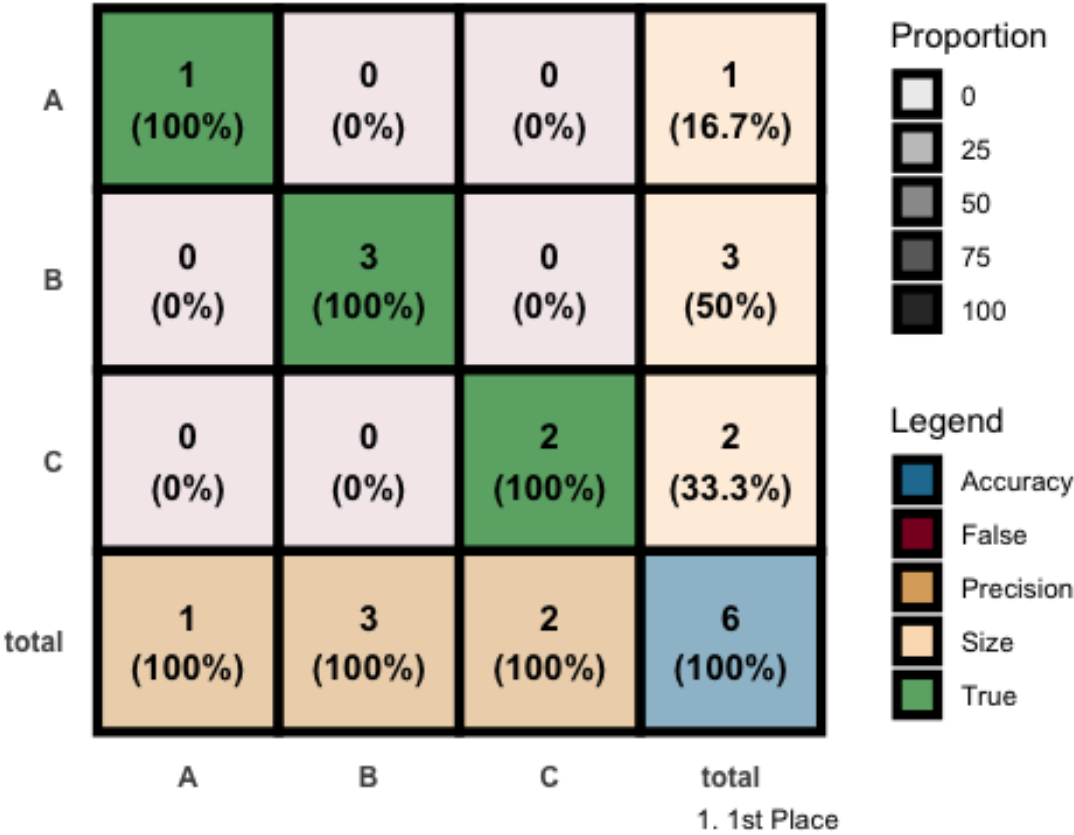
% probability



1. 1st Place

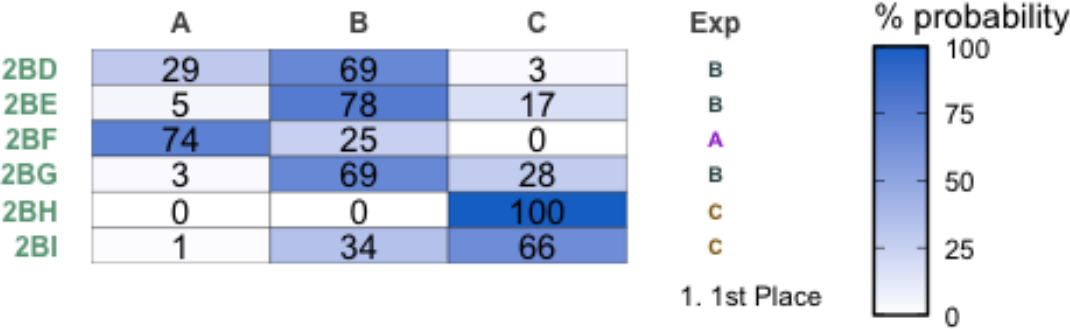
External Validation

Confusion Matrix



External Validation

Probability Heatmap



2nd Place

Table:

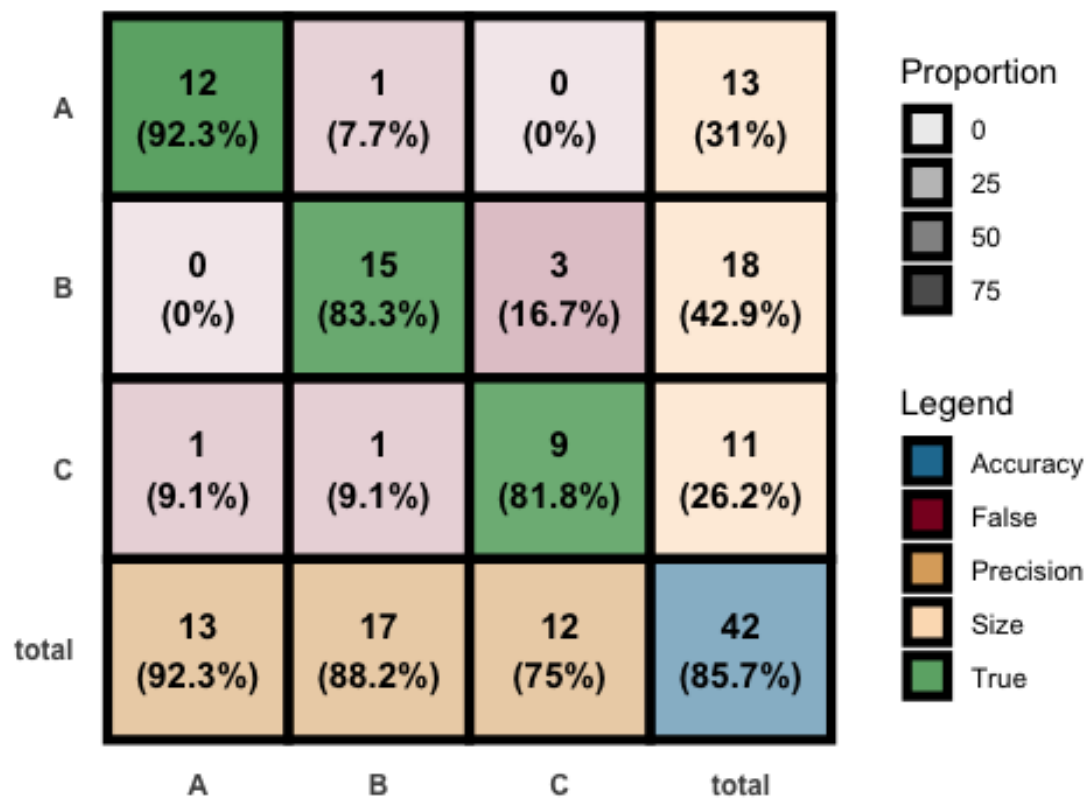
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.577

Model Coefficients

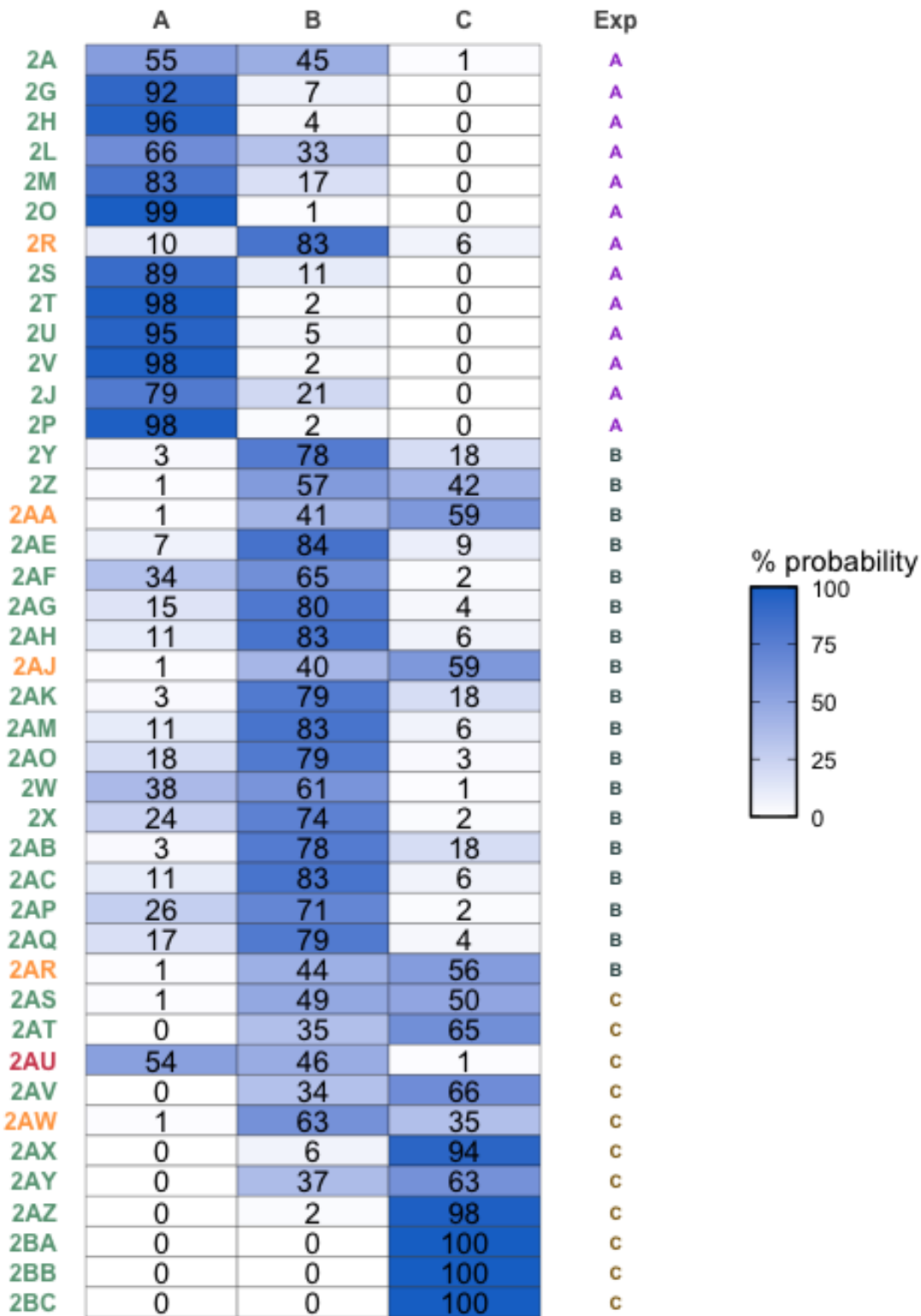
	x
-2-3-	-0.1316956
dip_y	-1.2386416
NPA_1	39.1279581
NPA_3	-227.3707397

Training Set Confusion Matrix



Training Set

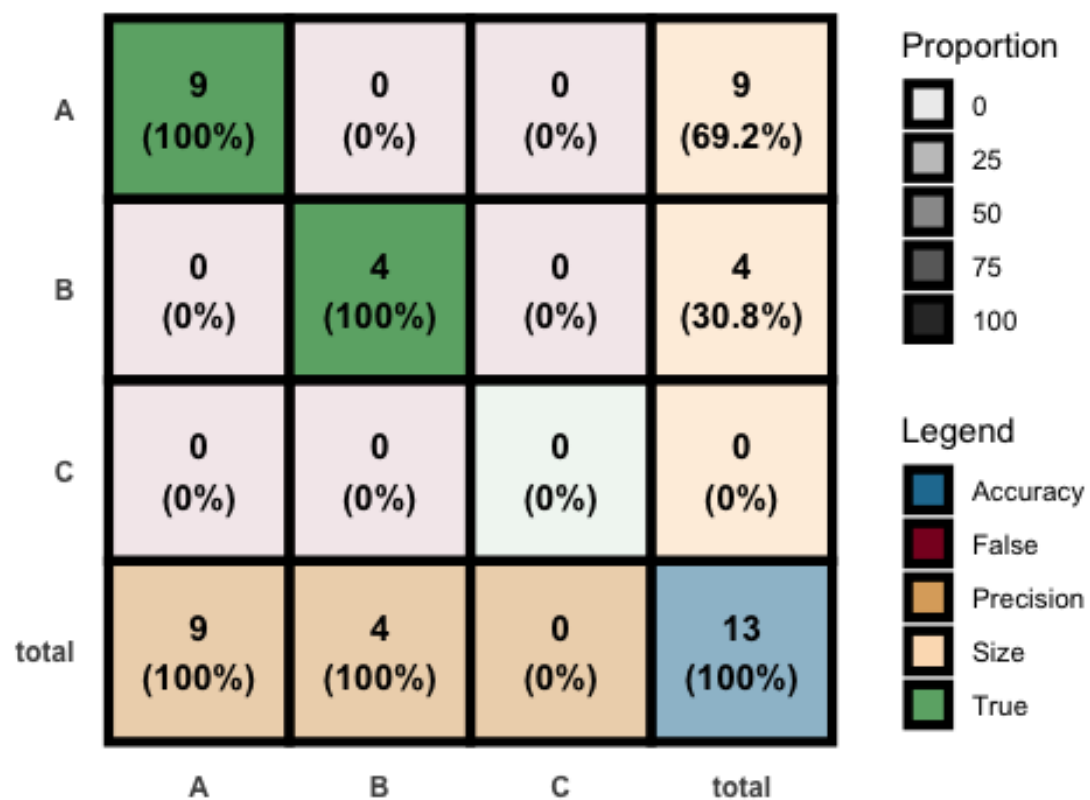
Probability Heatmap



2. 2nd Place

Test Set

Confusion Matrix



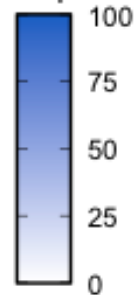
2. 2nd Place

Test Set

Probability Heatmap

	A	B	C	Exp
2B	83	17	0	A
2C	81	19	0	A
2D	100	0	0	A
2E	78	22	0	A
2F	79	21	0	A
2I	86	14	0	A
2K	100	0	0	A
2N	75	25	0	A
2Q	98	2	0	A
2AD	34	65	2	B
2AI	3	76	21	B
2AL	49	51	1	B
2AN	8	84	8	B

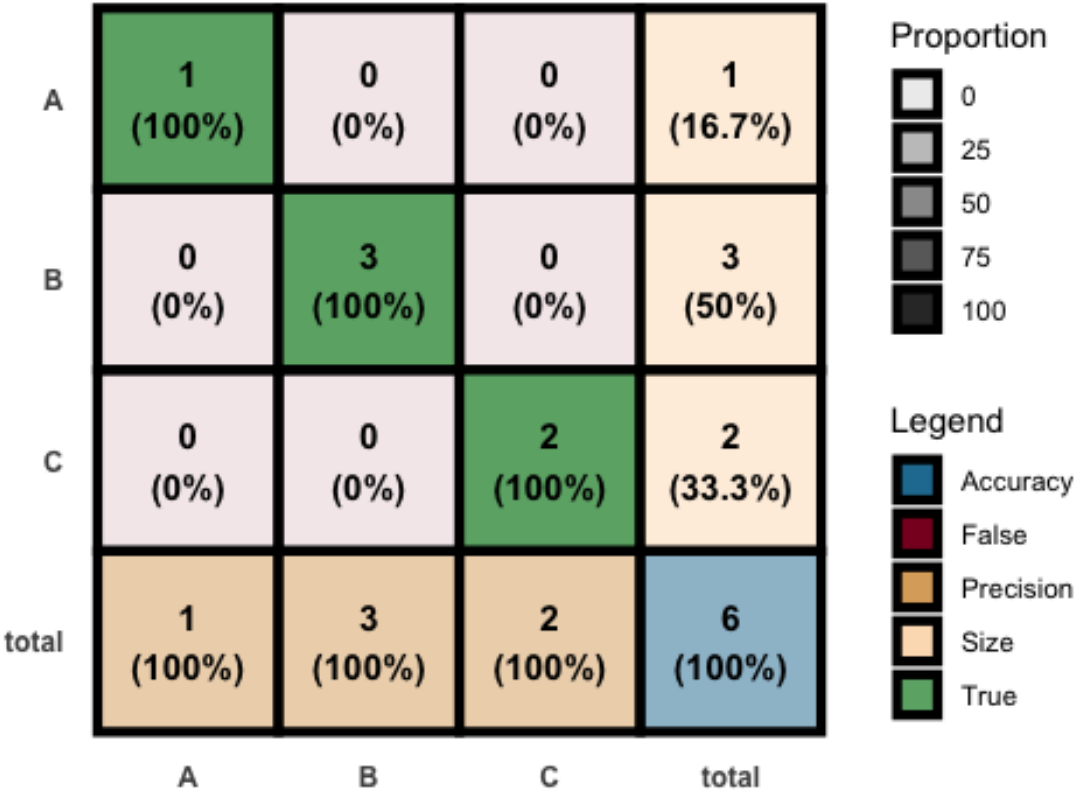
% probability



2. 2nd Place

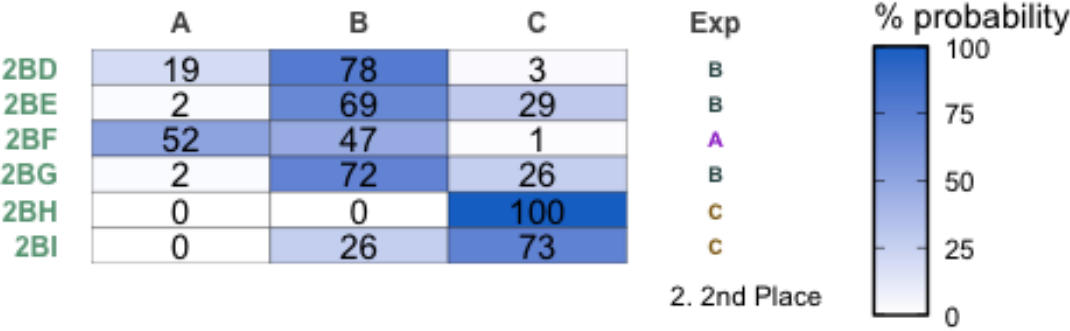
External Validation

Confusion Matrix



External Validation

Probability Heatmap



3rd Place

Table:

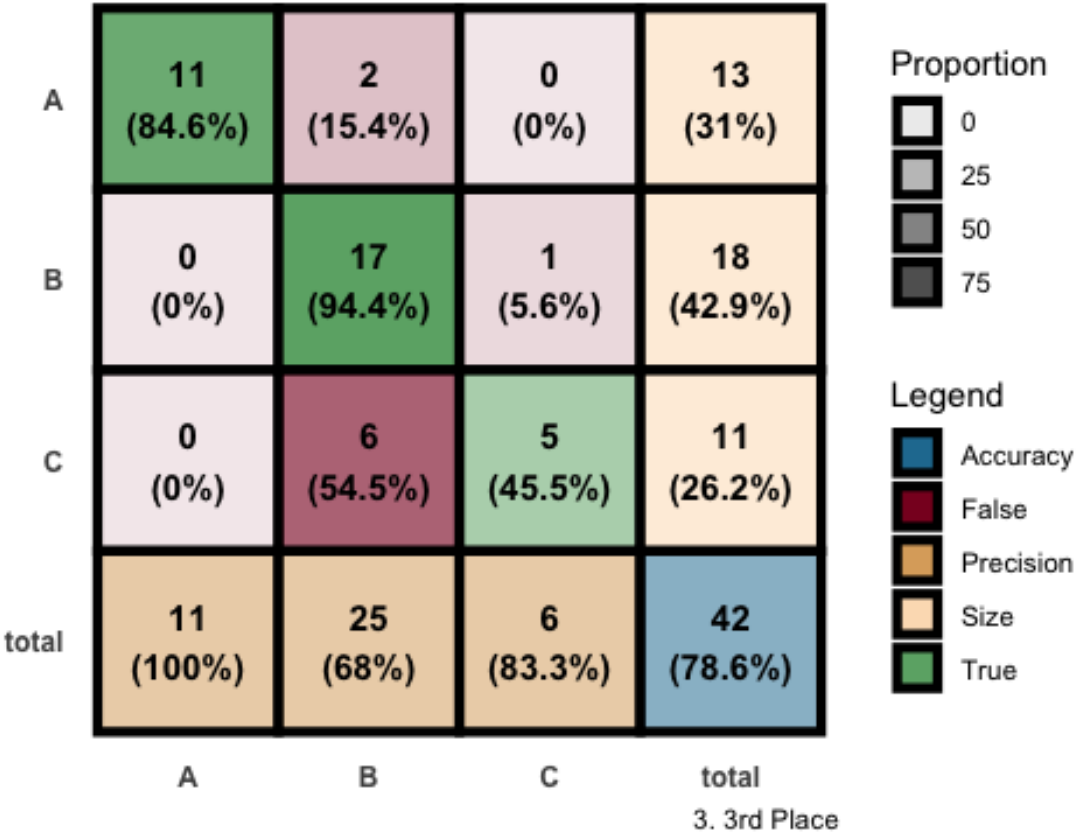
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
78.57%	0.57

Model Coefficients

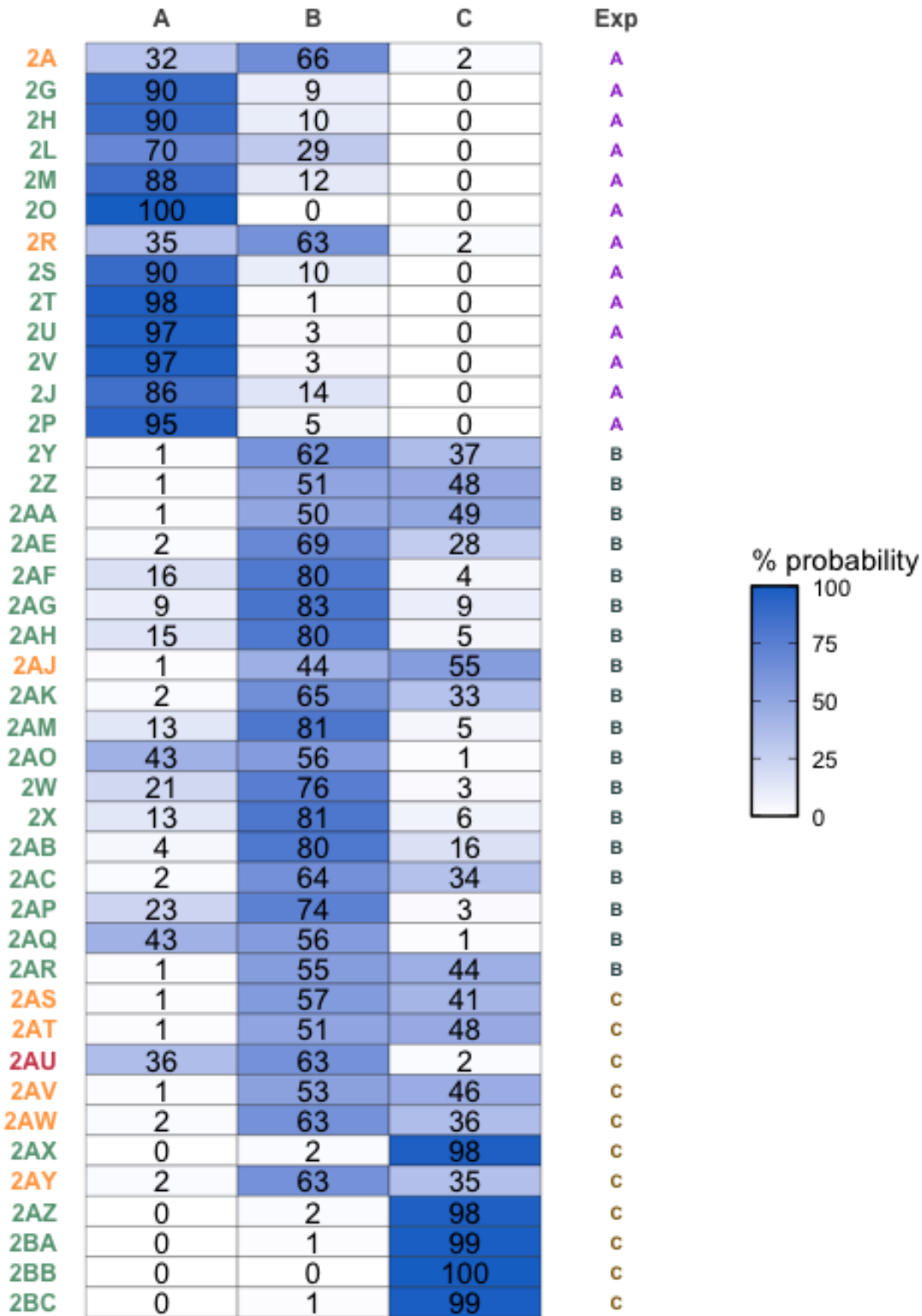
	x
dip_y	-1.9662387
NPA_3	-203.6733689
diff 1 2	-39.1768689
L	-0.4500423

Training Set
Confusion Matrix



Training Set

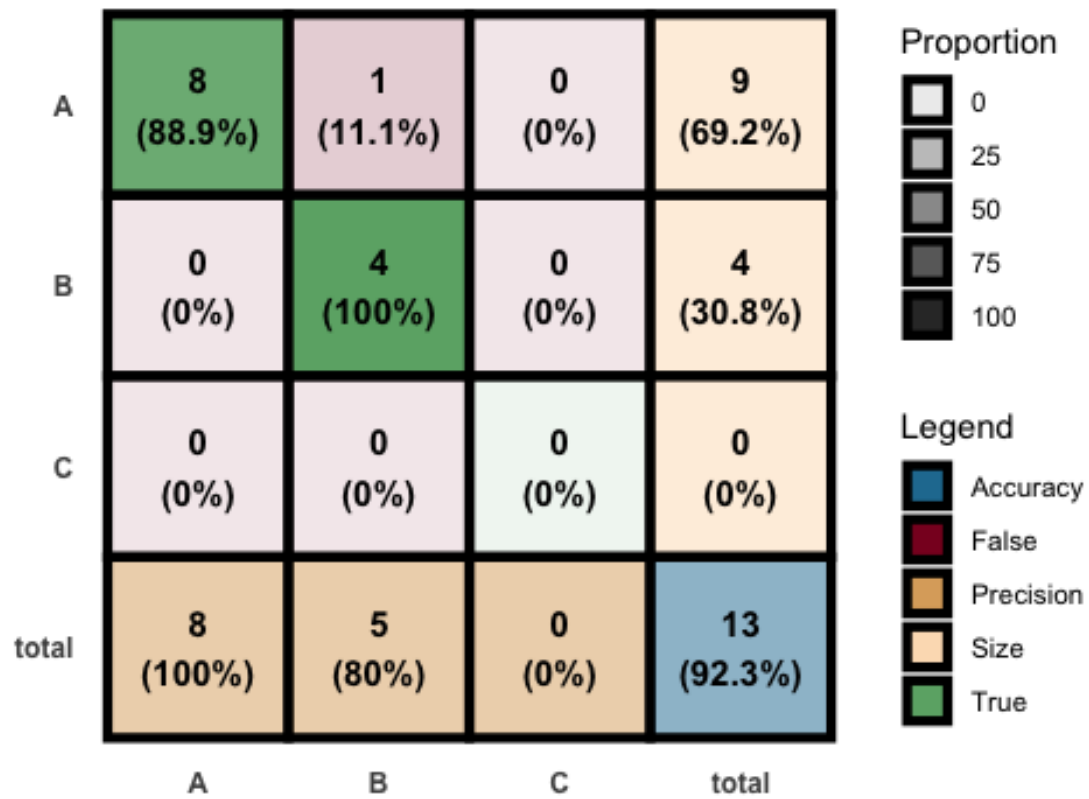
Probability Heatmap



3. 3rd Place

Test Set

Confusion Matrix



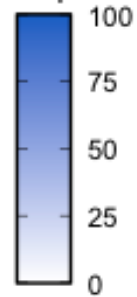
3. 3rd Place

Test Set

Probability Heatmap

	A	B	C	Exp
2B	57	42	1	A
2C	52	47	1	A
2D	99	1	0	A
2E	43	56	1	A
2F	54	45	1	A
2I	90	9	0	A
2K	100	0	0	A
2N	69	30	0	A
2Q	93	7	0	A
2AD	19	78	4	B
2AI	2	71	26	B
2AL	47	52	1	B
2AN	6	82	12	B

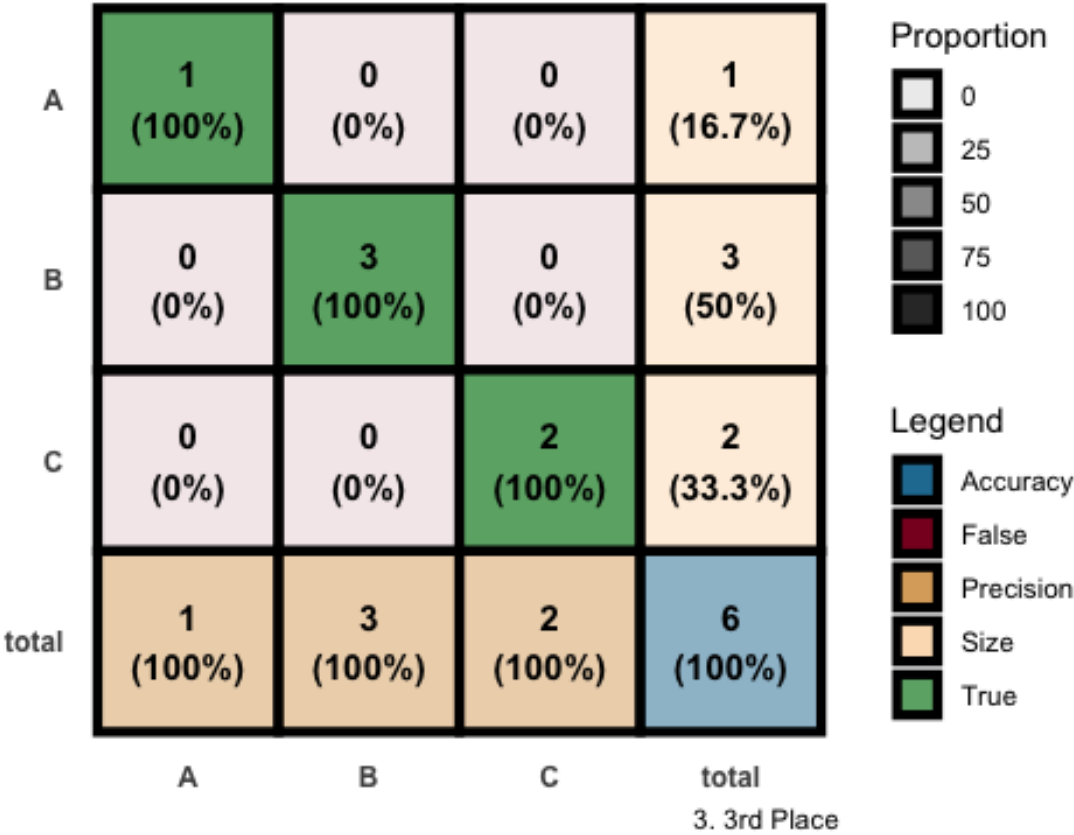
% probability



3. 3rd Place

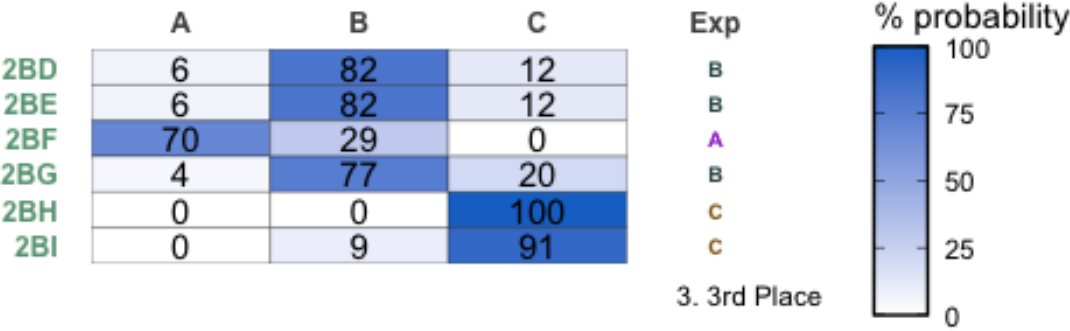
External Validation

Confusion Matrix



External Validation

Probability Heatmap



4th Place

Table:

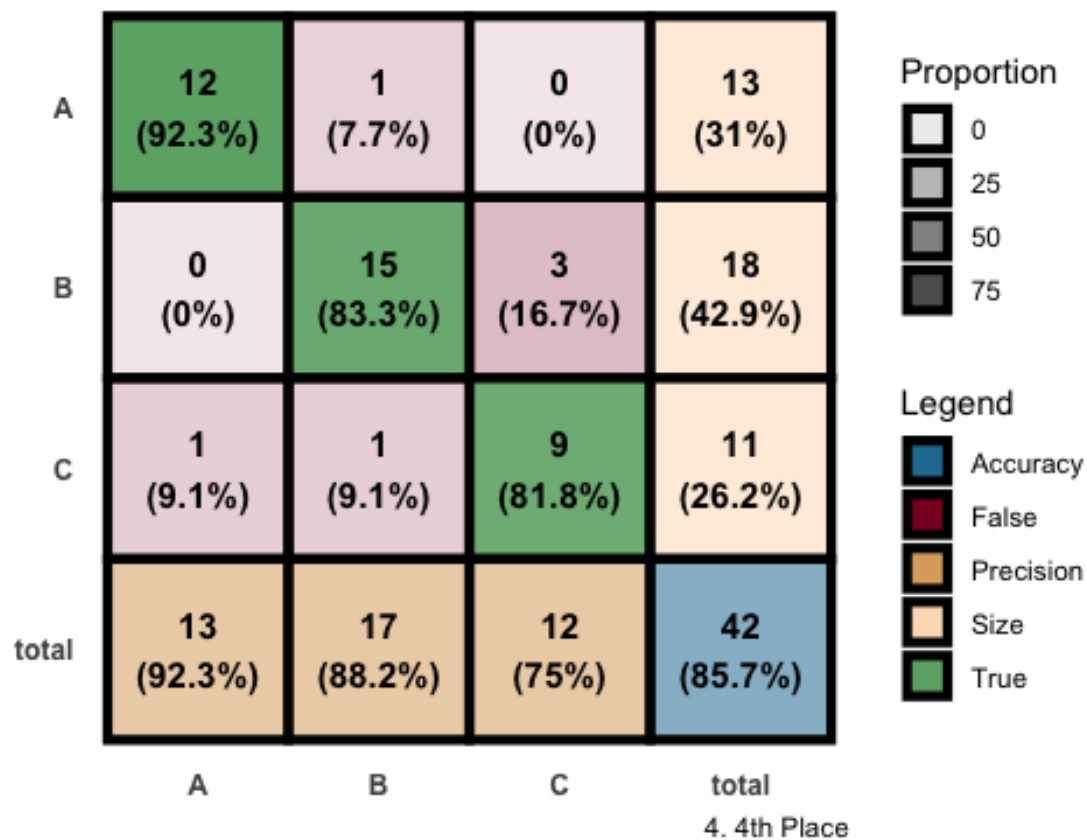
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.567

Model Coefficients

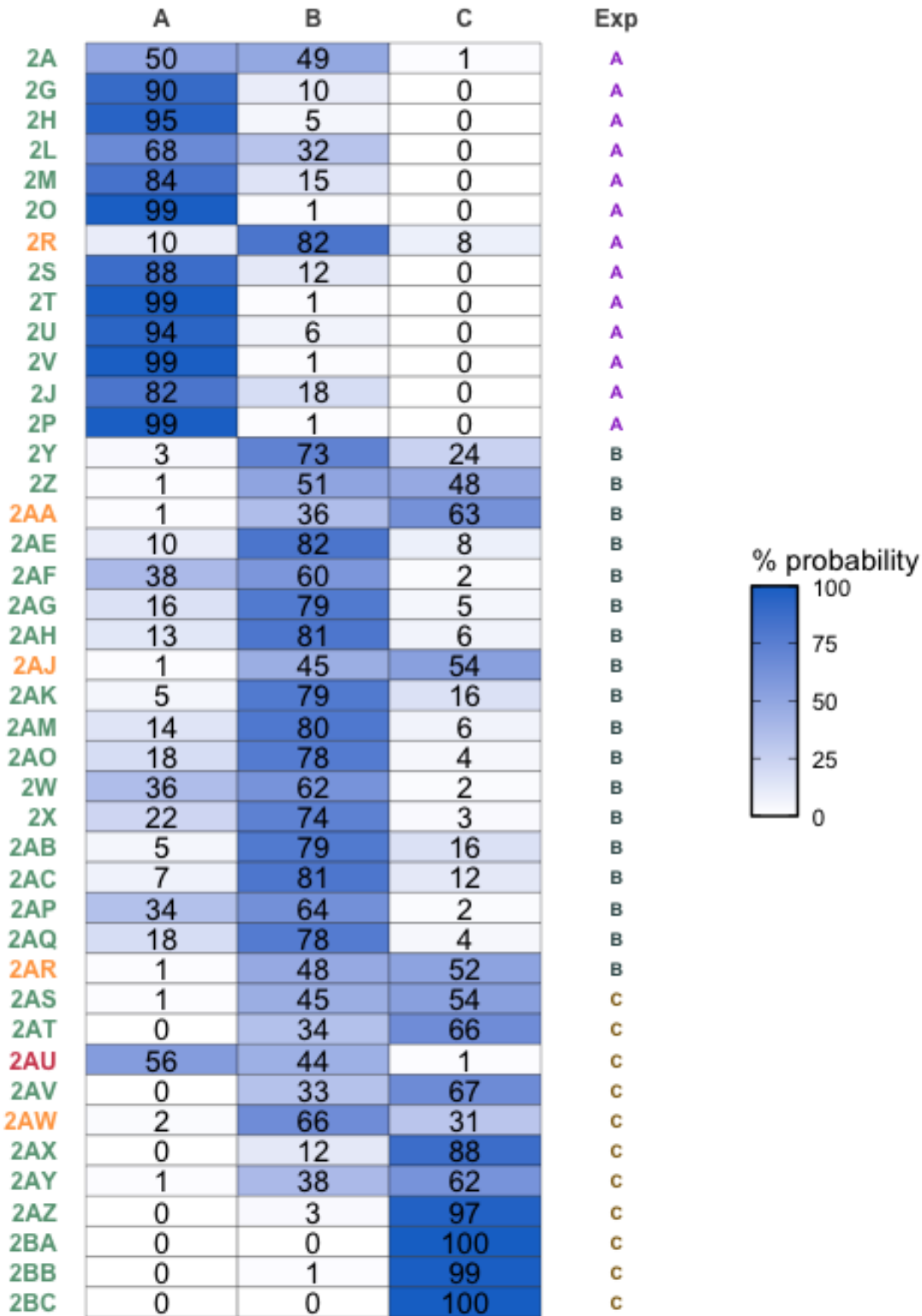
	x
dip_y	-1.212812
Dist(2, 7)	-205.902948
NPA_3	-298.027882
diff 1 2	-37.601198

Training Set Confusion Matrix



Training Set

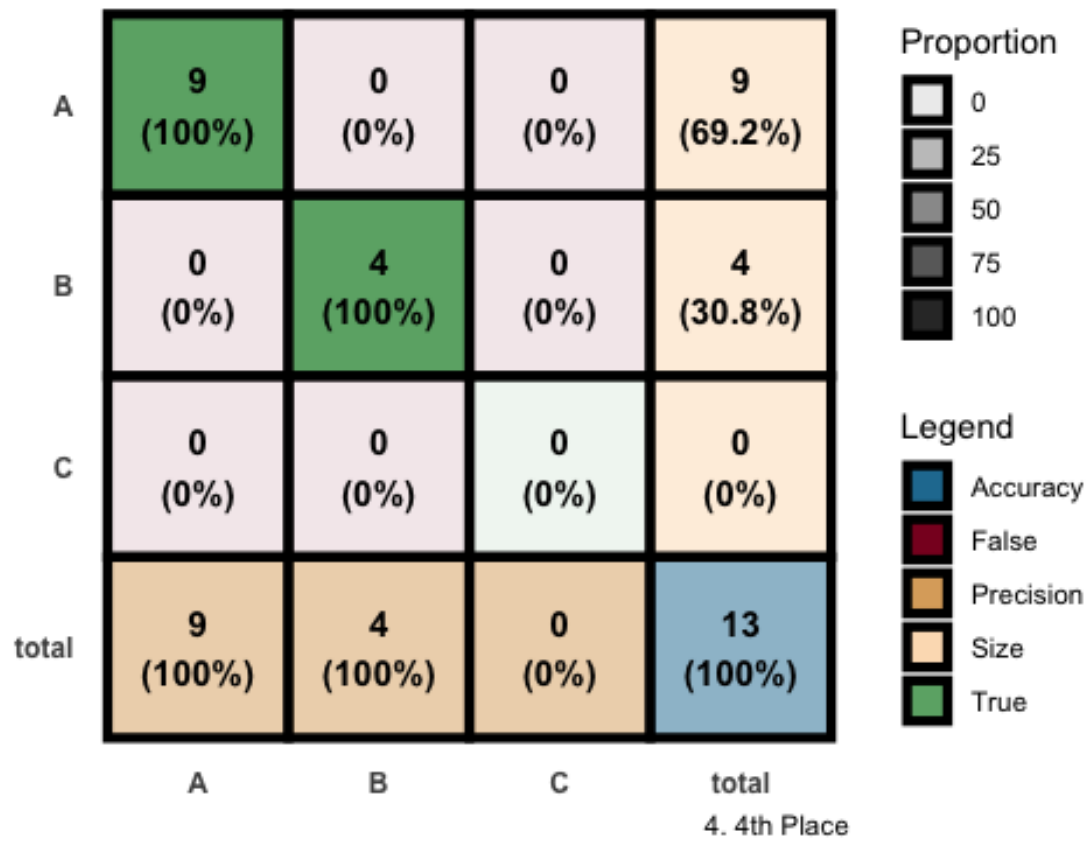
Probability Heatmap



4. 4th Place

Test Set

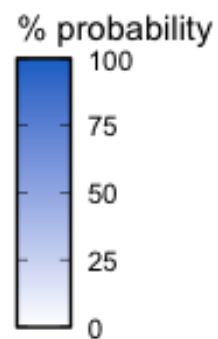
Confusion Matrix



Test Set

Probability Heatmap

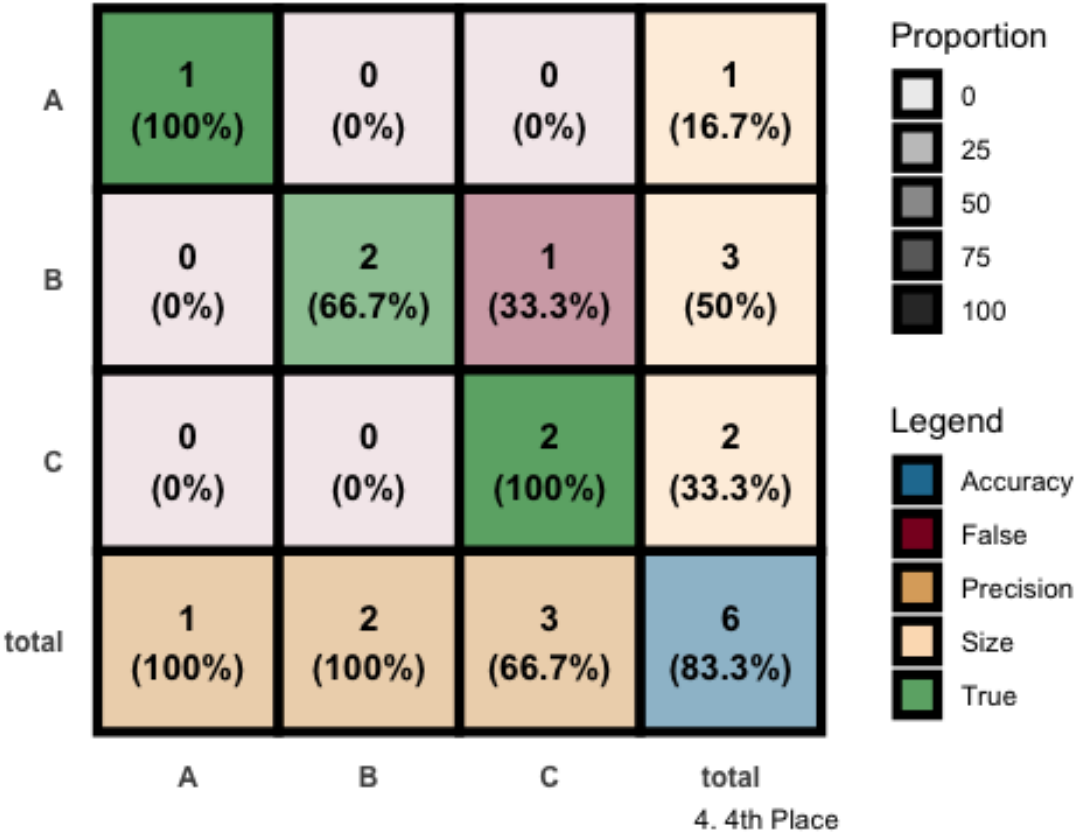
	A	B	C	Exp
2B	83	17	0	A
2C	81	19	0	A
2D	99	1	0	A
2E	75	24	0	A
2F	80	20	0	A
2I	87	13	0	A
2K	100	0	0	A
2N	74	25	0	A
2Q	98	2	0	A
2AD	29	69	2	B
2AI	4	78	18	B
2AL	47	52	1	B
2AN	9	82	9	B



4. 4th Place

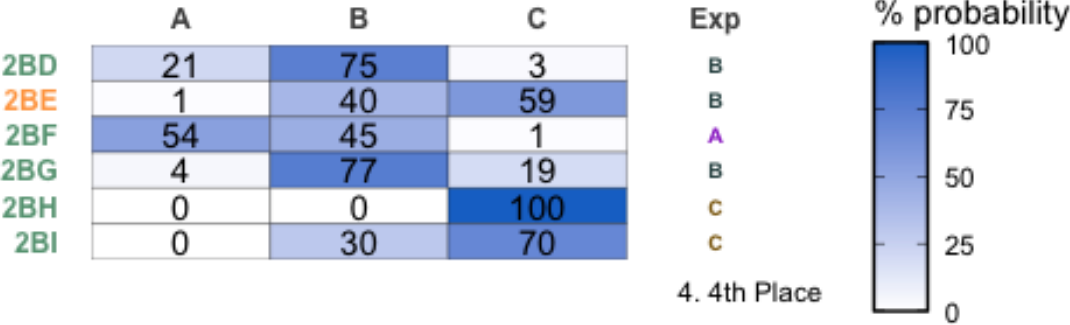
External Validation

Confusion Matrix



External Validation

Probability Heatmap



5th Place

Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
80.95%	0.557

Model Coefficients

	x
-2-3-	-0.2656046
dip_y	-1.5218006
NPA_1	37.0038197
NPA_2	58.0760567

Training Set Confusion Matrix

A	11 (84.6%)	2 (15.4%)	0 (0%)	13 (31%)
B	1 (5.6%)	15 (83.3%)	2 (11.1%)	18 (42.9%)
C	0 (0%)	3 (27.3%)	8 (72.7%)	11 (26.2%)
total	12 (91.7%)	20 (75%)	10 (80%)	42 (81%)
	A	B	C	total

5. 5th Place

Legend

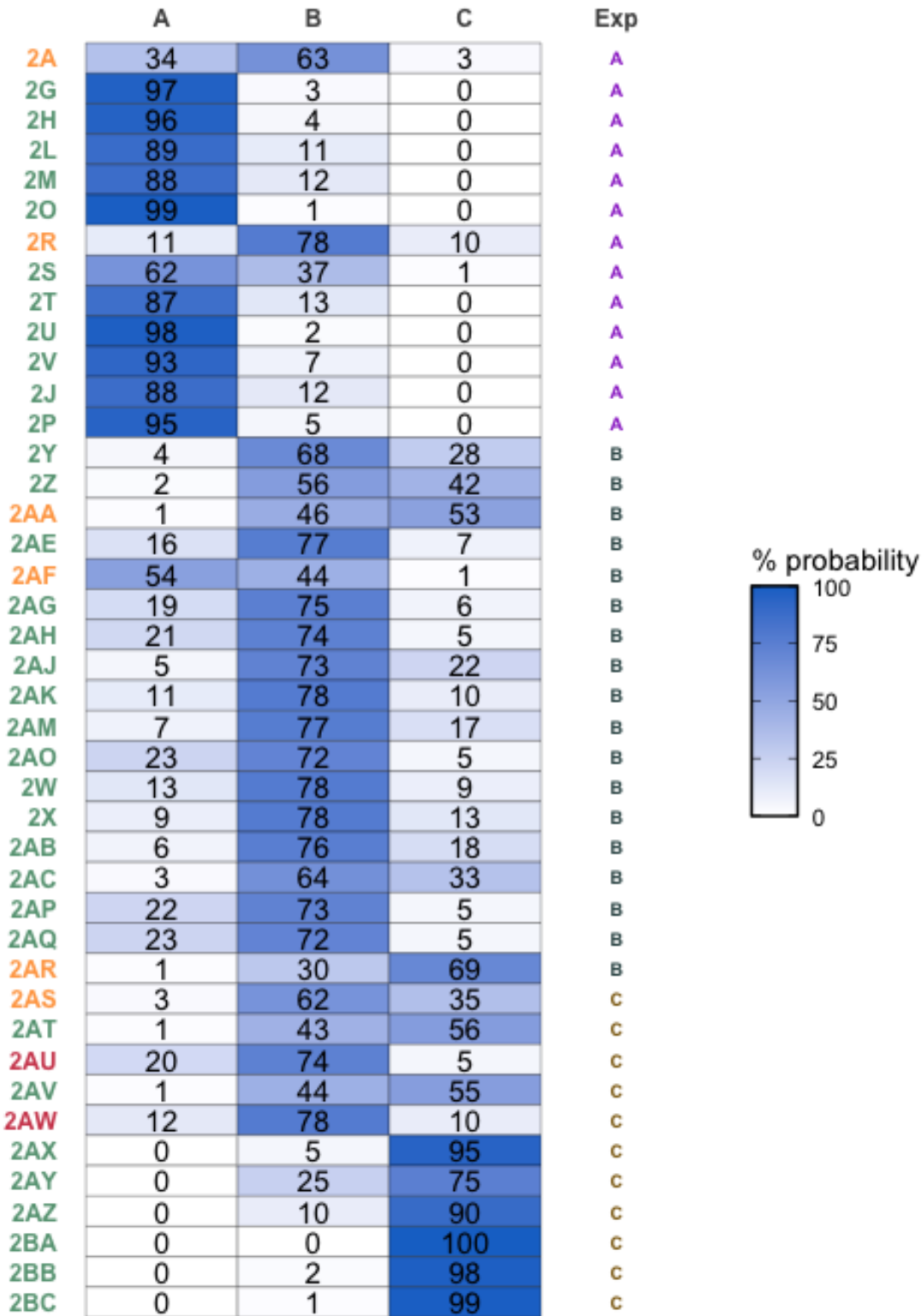
- Accuracy
- False
- Precision
- Size
- True

Proportion

- 0
- 20
- 40
- 60
- 80

Training Set

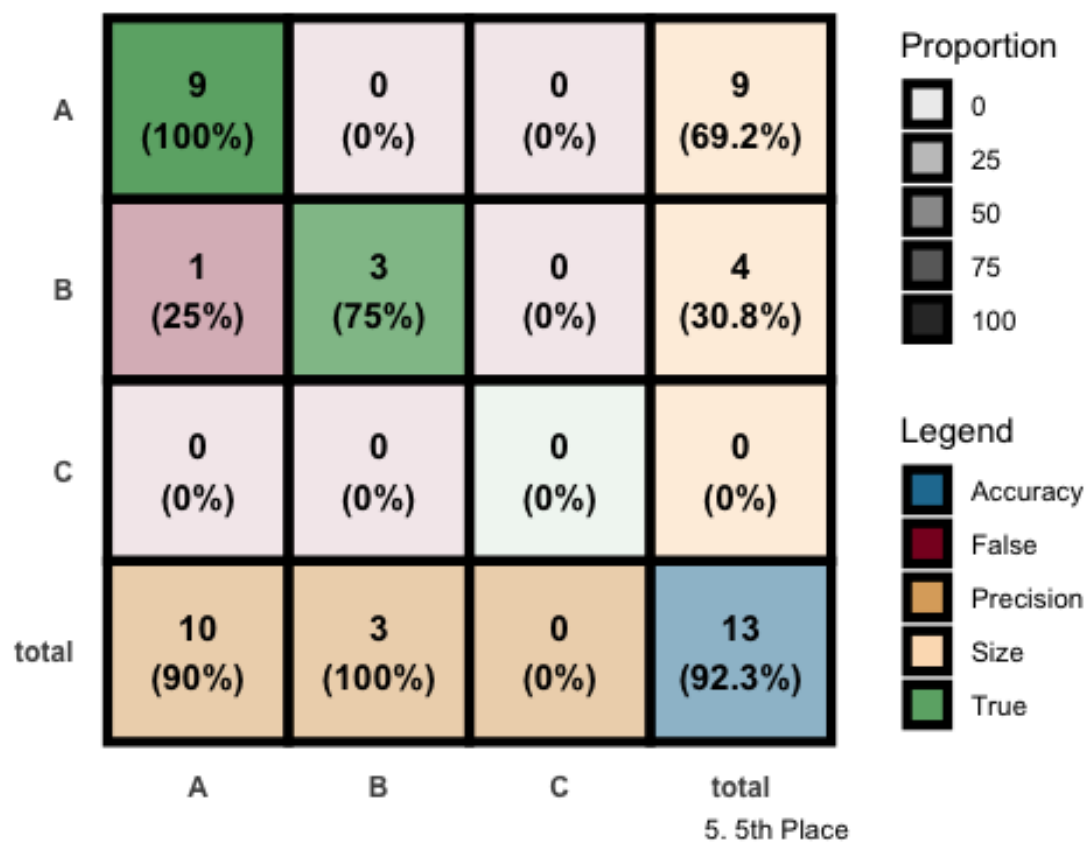
Probability Heatmap



5. 5th Place

Test Set

Confusion Matrix

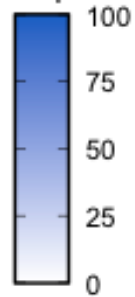


Test Set

Probability Heatmap

	A	B	C	Exp
2B	81	19	0	A
2C	81	19	0	A
2D	99	1	0	A
2E	79	21	0	A
2F	76	24	0	A
2I	92	8	0	A
2K	100	0	0	A
2N	75	25	0	A
2Q	95	5	0	A
2AD	15	78	8	B
2AI	9	78	13	B
2AL	61	38	1	B
2AN	9	78	13	B

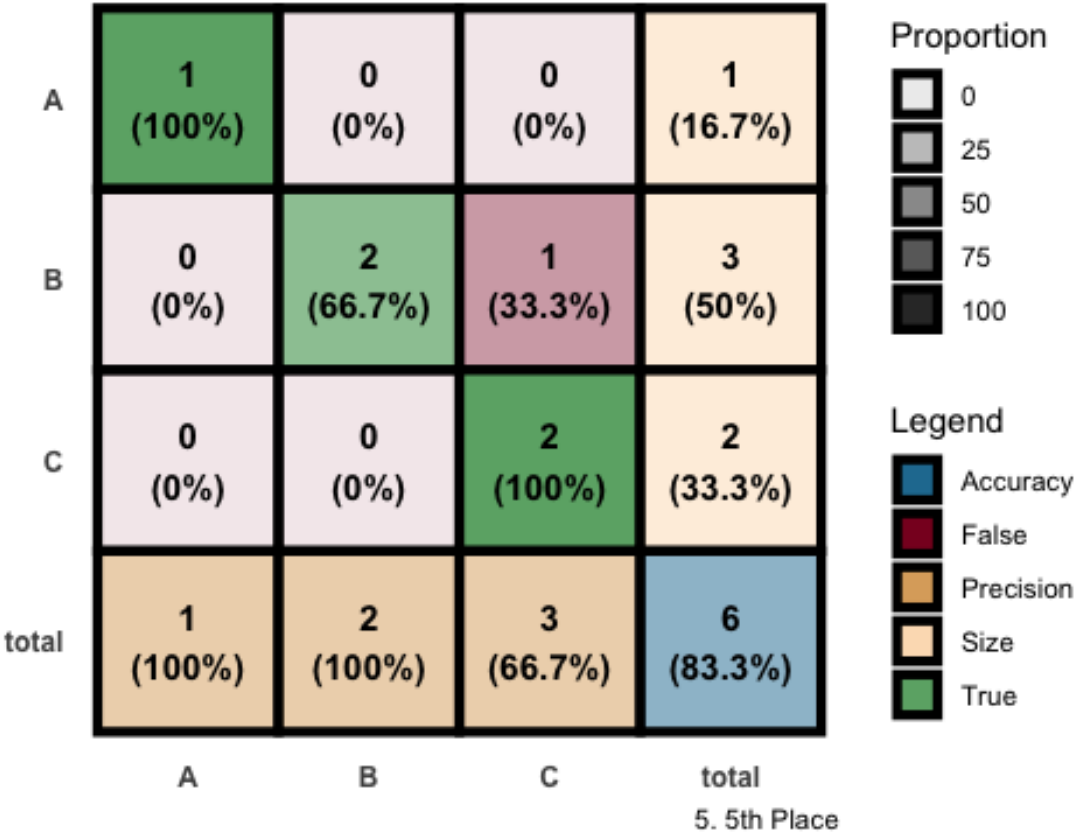
% probability



5. 5th Place

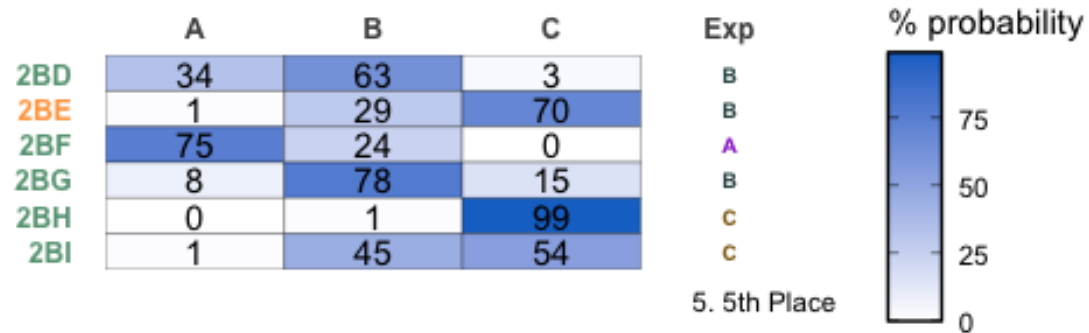
External Validation

Confusion Matrix



External Validation

Probability Heatmap



6th Place

Table:

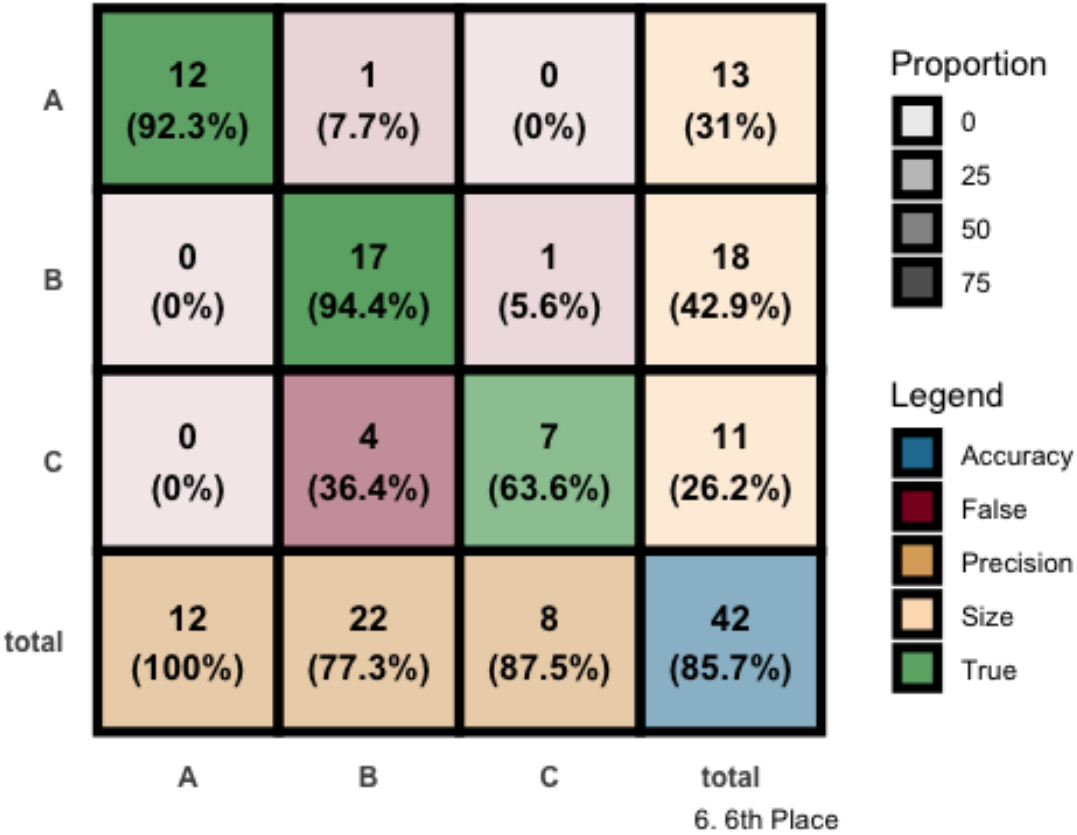
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.555

Model Coefficients

	x
-2-3-	-0.298887
dip_y	-1.568930
Dist(2, 7)	243.537174
diff 1 2	-42.690753

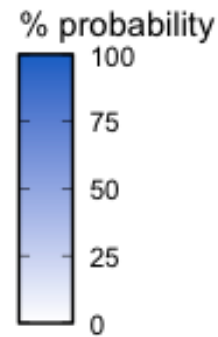
Training Set
Confusion Matrix



Training Set

Probability Heatmap

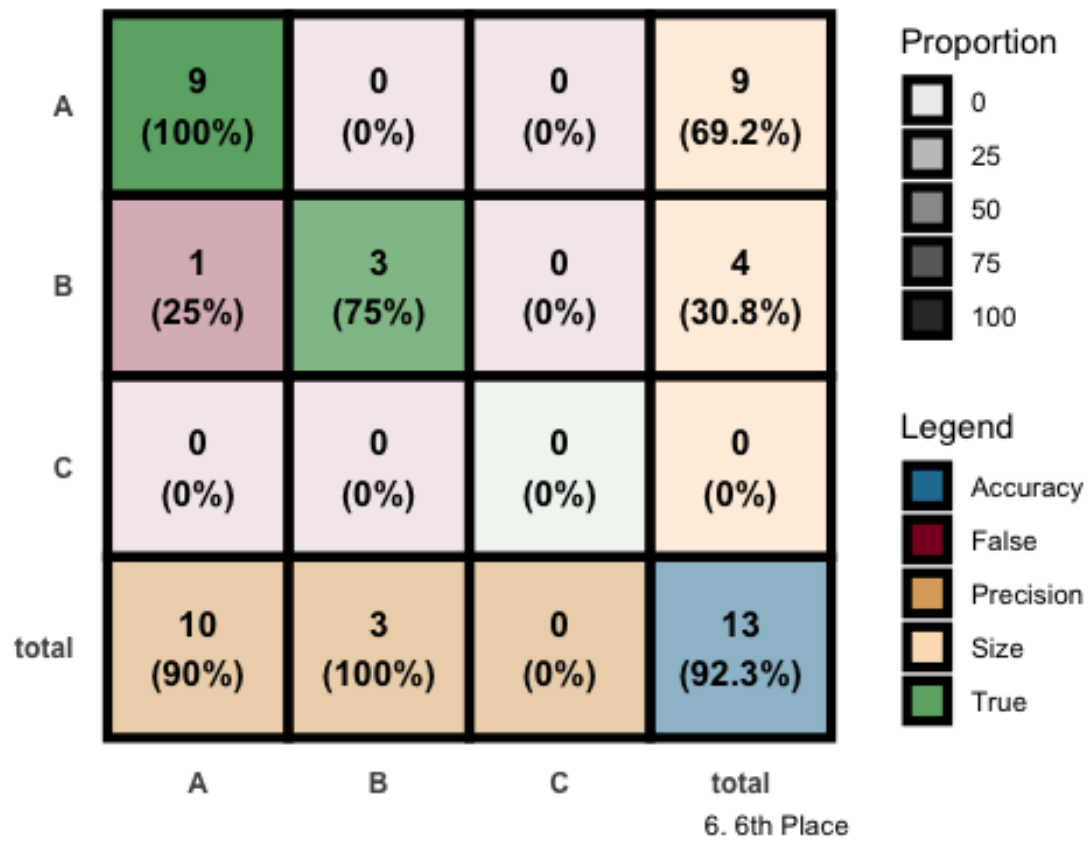
	A	B	C	Exp
2A	53	46	1	A
2G	98	2	0	A
2H	96	3	0	A
2L	89	11	0	A
2M	88	12	0	A
2O	99	1	0	A
2R	12	77	10	A
2S	69	30	1	A
2T	64	35	1	A
2U	99	1	0	A
2V	76	24	1	A
2J	87	12	0	A
2P	97	3	0	A
2Y	3	62	35	B
2Z	3	64	33	B
2AA	2	56	42	B
2AE	13	77	10	B
2AF	49	49	2	B
2AG	17	75	7	B
2AH	18	75	7	B
2AJ	3	64	33	B
2AK	9	77	14	B
2AM	11	77	12	B
2AO	22	73	6	B
2W	27	69	4	B
2X	16	76	8	B
2AB	5	71	24	B
2AC	4	68	28	B
2AP	25	71	5	B
2AQ	21	73	6	B
2AR	1	31	68	B
2AS	1	45	53	C
2AT	2	51	47	C
2AU	39	59	3	C
2AV	2	54	44	C
2AW	9	77	14	C
2AX	0	3	97	C
2AY	1	29	70	C
2AZ	0	4	96	C
2BA	0	0	100	C
2BB	0	0	100	C
2BC	0	0	100	C



6. 6th Place

Test Set

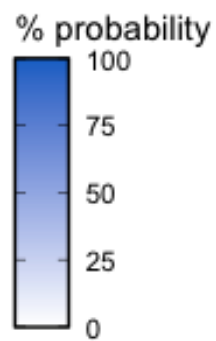
Confusion Matrix



Test Set

Probability Heatmap

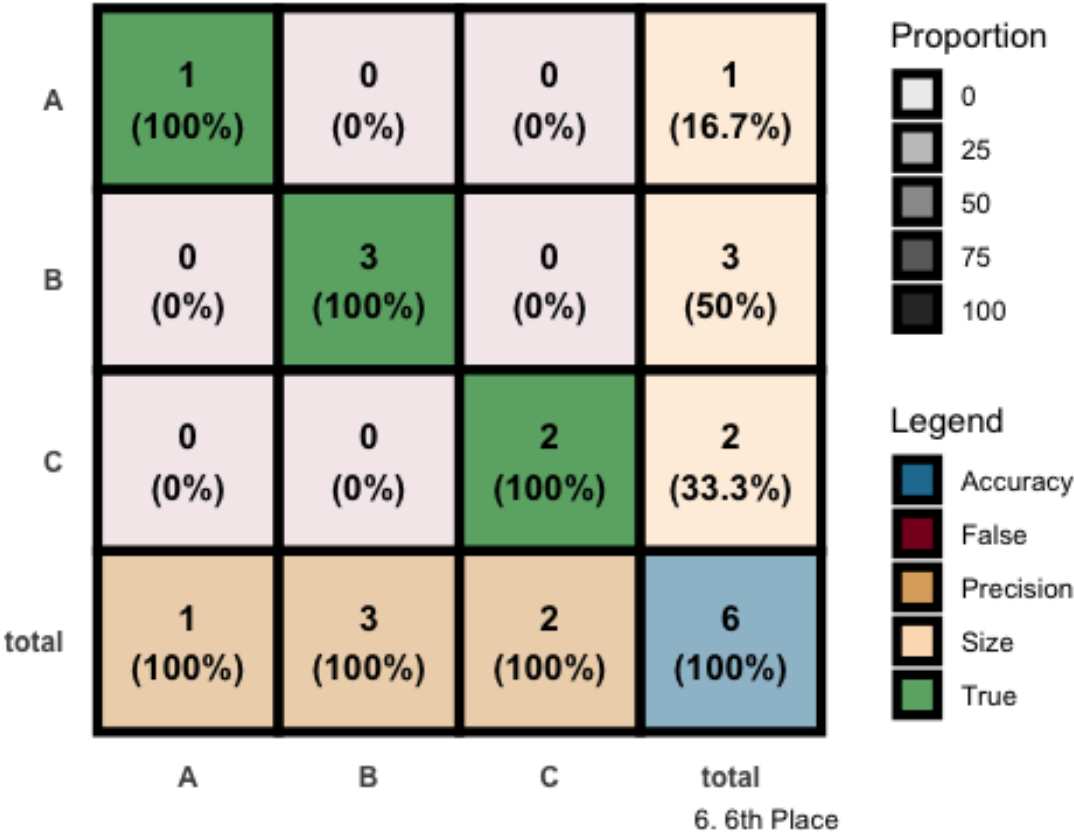
	A	B	C	Exp
2B	83	16	0	A
2C	82	18	0	A
2D	100	0	0	A
2E	78	21	0	A
2F	78	21	0	A
2I	92	8	0	A
2K	100	0	0	A
2N	71	28	1	A
2Q	96	4	0	A
2AD	27	68	4	B
2AI	7	75	18	B
2AL	64	36	1	B
2AN	10	77	13	B



6. 6th Place

External Validation

Confusion Matrix



External Validation

Probability Heatmap

	A	B	C
2BD	31	66	4
2BE	7	75	18
2BF	75	25	1
2BG	4	69	26
2BH	0	0	100
2BI	1	36	63

Exp

B

B

A

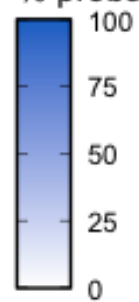
B

C

C

6. 6th Place

% probability



7th Place

Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
83.33%	0.555

Model Coefficients

	x
-2-7-	0.0196432
dip_y	-1.2374931
NPA_1	34.2137827
NPA_3	-247.0244428

Training Set Confusion Matrix

A	11 (84.6%)	2 (15.4%)	0 (0%)	13 (31%)
B	0 (0%)	16 (88.9%)	2 (11.1%)	18 (42.9%)
C	0 (0%)	3 (27.3%)	8 (72.7%)	11 (26.2%)
total	11 (100%)	21 (76.2%)	10 (80%)	42 (83.3%)
	A	B	C	total

7. 7th Place

Legend

- Accuracy
- False
- Precision
- Size
- True

Proportion

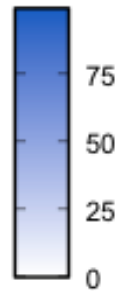
- 0
- 20
- 40
- 60
- 80

Training Set

Probability Heatmap

	A	B	C	Exp
2A	47	51	2	A
2G	88	12	0	A
2H	94	6	0	A
2L	71	29	1	A
2M	81	18	0	A
2O	99	1	0	A
2R	8	78	14	A
2S	85	14	0	A
2T	99	1	0	A
2U	93	7	0	A
2V	99	1	0	A
2J	81	18	0	A
2P	97	3	0	A
2Y	4	68	28	B
2Z	2	51	47	B
2AA	1	37	62	B
2AE	10	78	11	B
2AF	40	58	2	B
2AG	13	78	9	B
2AH	13	78	9	B
2AJ	1	47	52	B
2AK	6	75	19	B
2AM	15	77	8	B
2AO	16	77	7	B
2W	31	66	3	B
2X	19	75	6	B
2AB	8	77	16	B
2AC	6	74	20	B
2AP	35	63	3	B
2AQ	16	77	7	B
2AR	2	49	49	B
2AS	2	53	46	C
2AT	1	39	60	C
2AU	48	51	2	C
2AV	1	38	61	C
2AW	4	68	28	C
2AX	0	13	87	C
2AY	1	41	58	C
2AZ	0	4	95	C
2BA	0	2	98	C
2BB	0	7	93	C
2BC	0	2	98	C

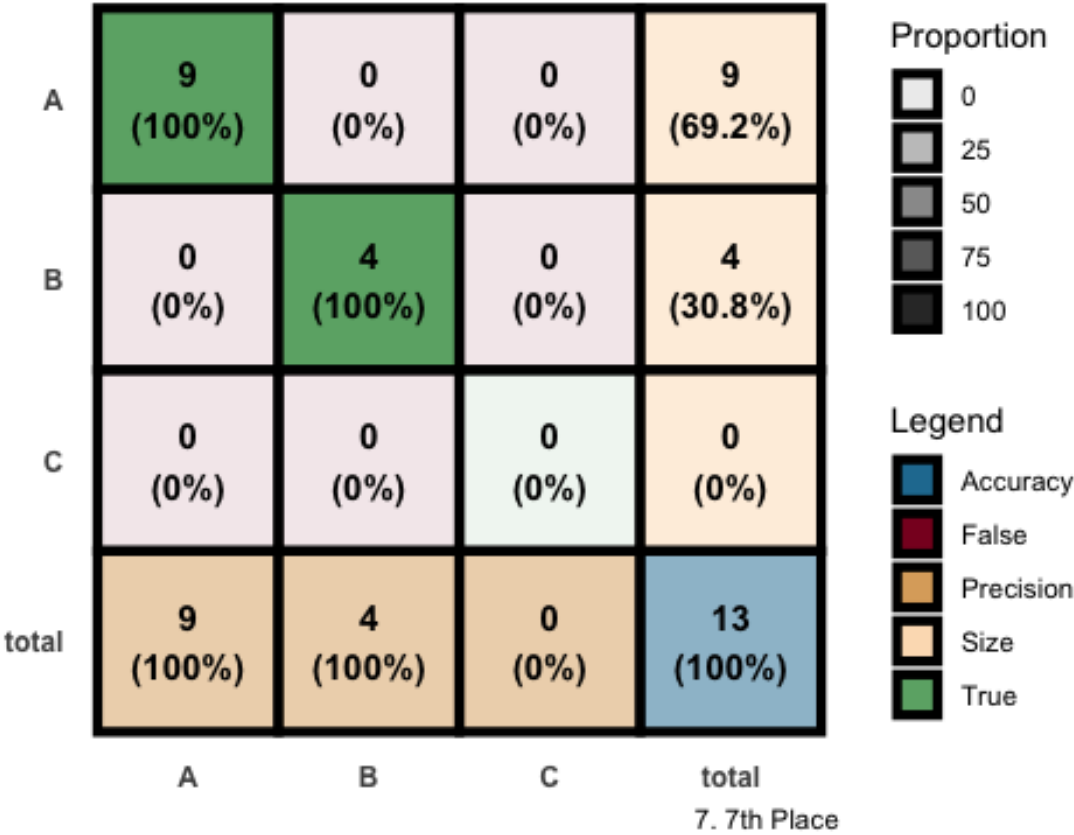
% probability



7. 7th Place

Test Set

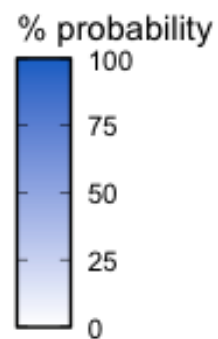
Confusion Matrix



Test Set

Probability Heatmap

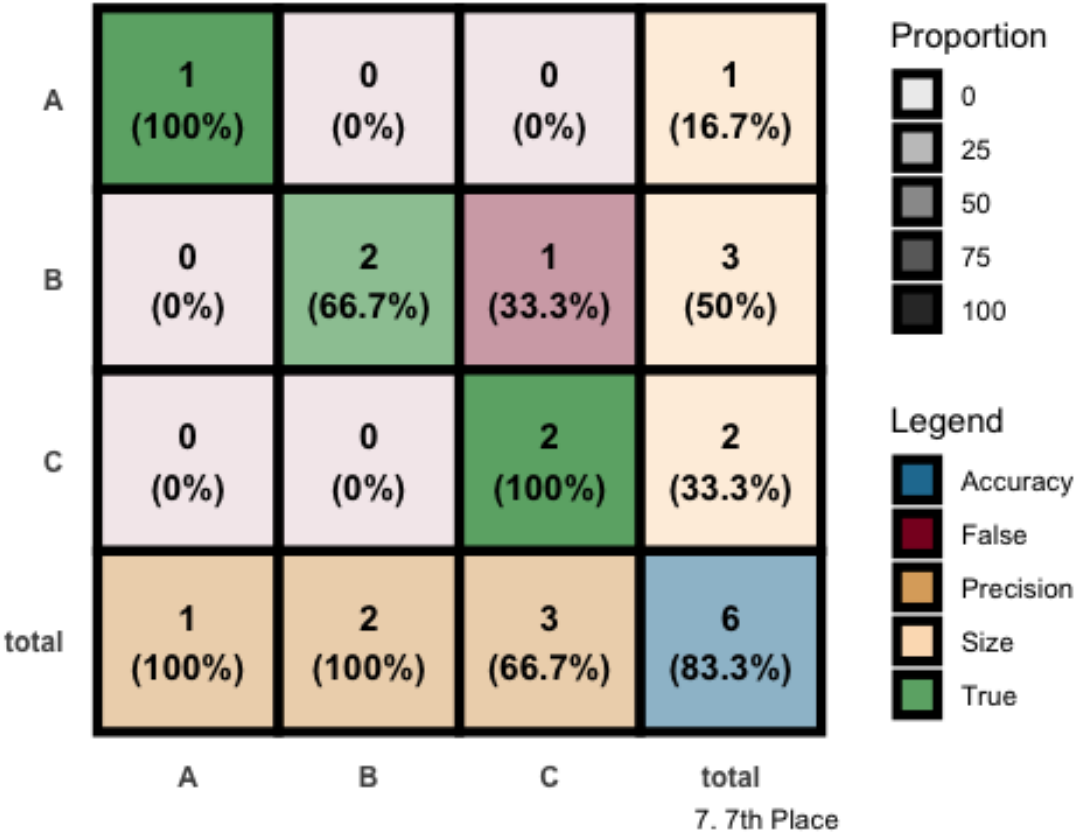
	A	B	C	Exp
2B	78	21	0	A
2C	75	24	0	A
2D	99	1	0	A
2E	70	30	1	A
2F	74	26	1	A
2I	85	15	0	A
2K	100	0	0	A
2N	69	31	1	A
2Q	96	4	0	A
2AD	28	68	4	B
2AI	5	73	22	B
2AL	45	54	2	B
2AN	7	77	16	B



7. 7th Place

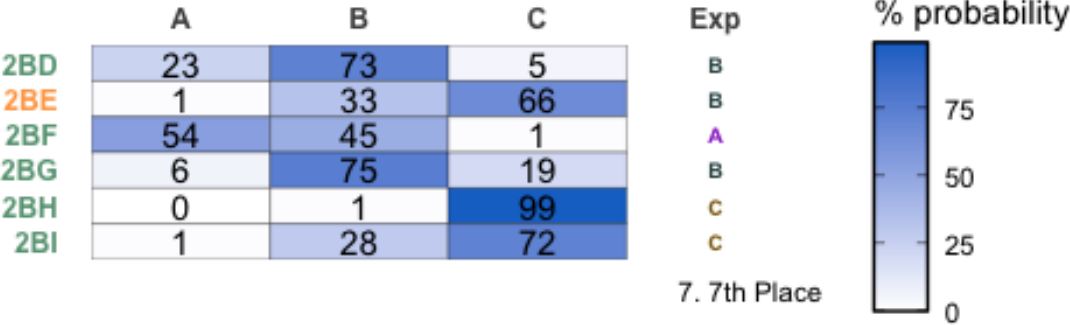
External Validation

Confusion Matrix



External Validation

Probability Heatmap



8th Place

Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
83.33%	0.553

Model Coefficients

	x
dip_y	-1.267705
NPA_1	37.769555
NPA_3	-256.395854
diff 2 7	-60.498122

Training Set Confusion Matrix

A	11 (84.6%)	2 (15.4%)	0 (0%)	13 (31%)
B	0 (0%)	16 (88.9%)	2 (11.1%)	18 (42.9%)
C	0 (0%)	3 (27.3%)	8 (72.7%)	11 (26.2%)
total	11 (100%)	21 (76.2%)	10 (80%)	42 (83.3%)
	A	B	C	total

Legend

- Accuracy
- False
- Precision
- Size
- True

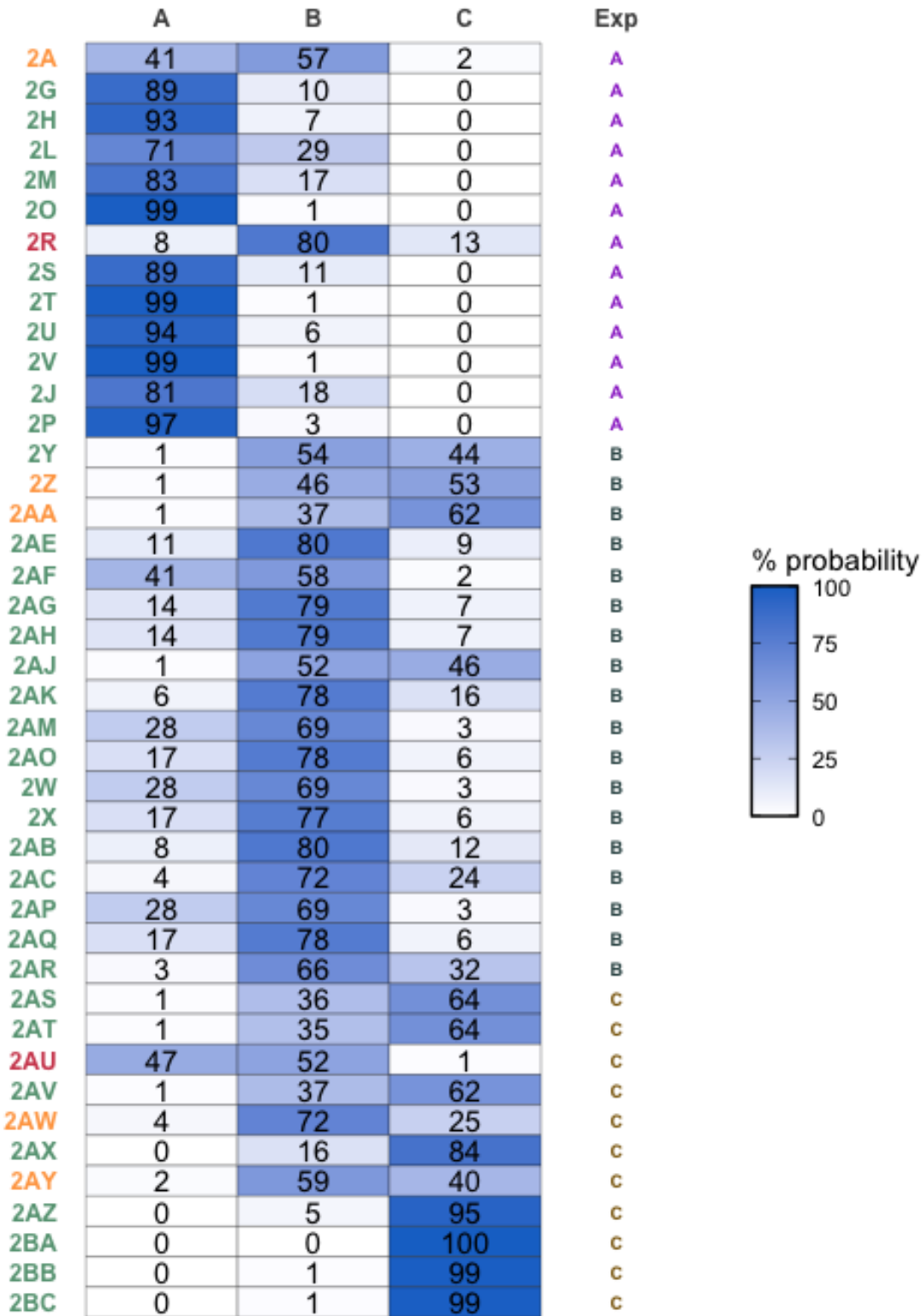
Proportion

- 0
- 20
- 40
- 60
- 80

8. 8th Place

Training Set

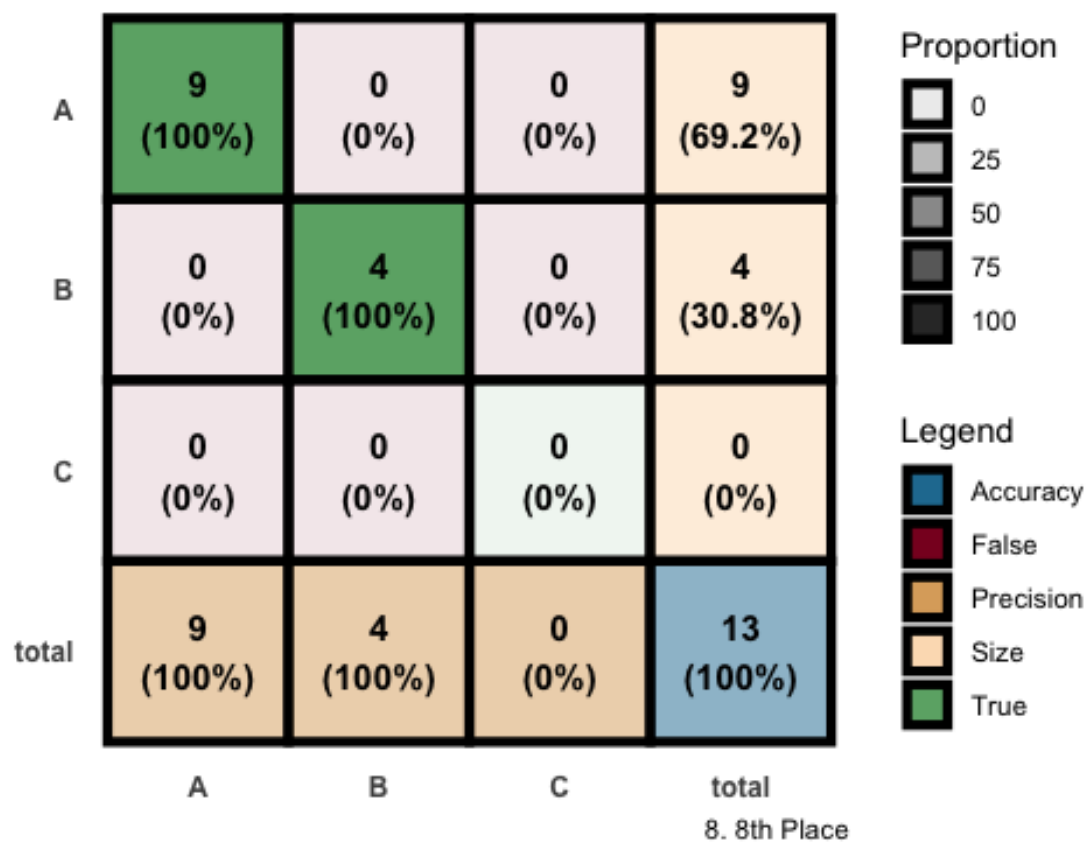
Probability Heatmap



8. 8th Place

Test Set

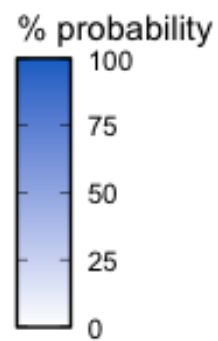
Confusion Matrix



Test Set

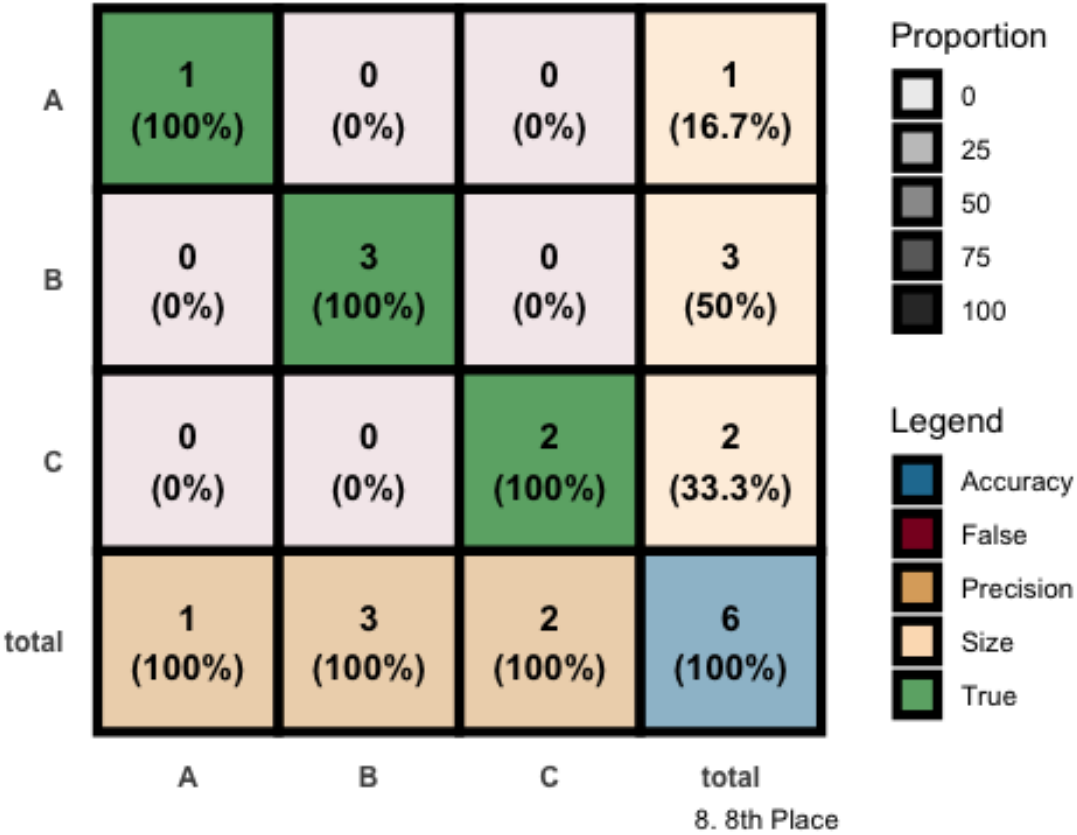
Probability Heatmap

	A	B	C	Exp
2B	78	22	0	A
2C	75	25	0	A
2D	99	1	0	A
2E	69	30	1	A
2F	74	25	0	A
2I	86	14	0	A
2K	100	0	0	A
2N	67	33	1	A
2Q	96	4	0	A
2AD	21	75	4	B
2AI	5	77	18	B
2AL	44	55	2	B
2AN	7	79	14	B



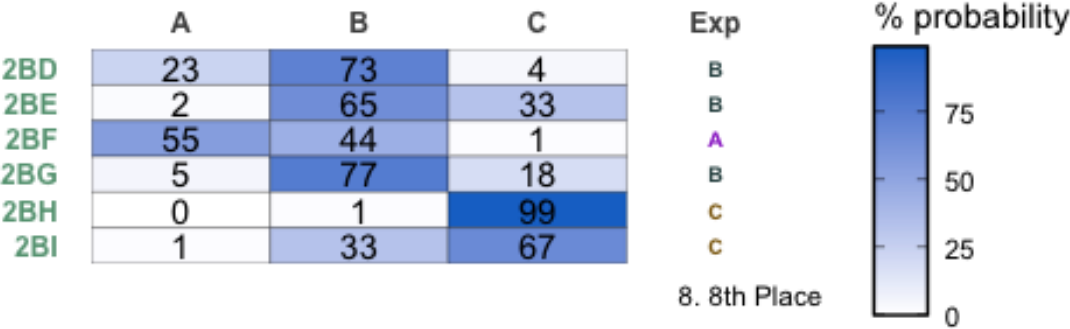
8. 8th Place

External Validation
Confusion Matrix



External Validation

Probability Heatmap



9th Place

Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
83.33%	0.551

Model Coefficients

	x
dip_y	-1.311817
NPA_3	-250.122843
NPA_7	57.898207
diff 1 2	-37.003103

Training Set Confusion Matrix

A	11 (84.6%)	2 (15.4%)	0 (0%)	13 (31%)
B	0 (0%)	16 (88.9%)	2 (11.1%)	18 (42.9%)
C	0 (0%)	3 (27.3%)	8 (72.7%)	11 (26.2%)
total	11 (100%)	21 (76.2%)	10 (80%)	42 (83.3%)
	A	B	C	total

Legend

- Accuracy
- False
- Precision
- Size
- True

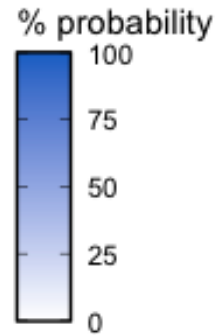
Proportion

- 0
- 20
- 40
- 60
- 80

Training Set

Probability Heatmap

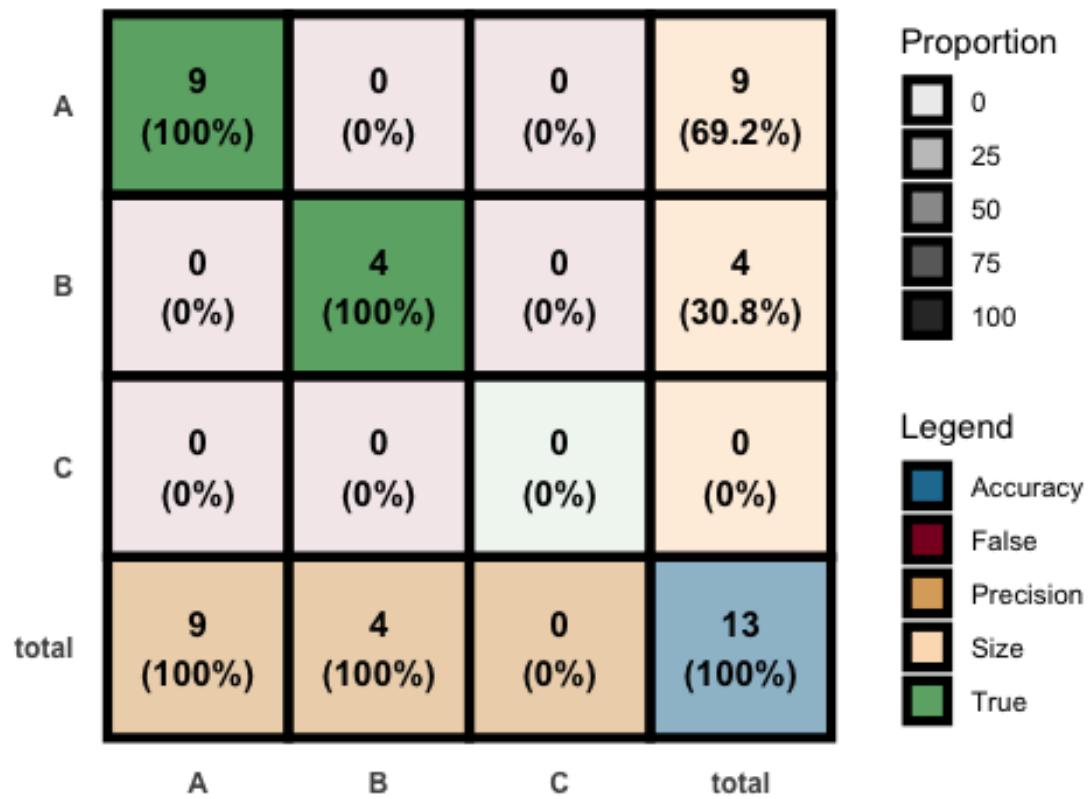
	A	B	C	Exp
2A	42	56	2	A
2G	89	10	0	A
2H	93	7	0	A
2L	71	29	1	A
2M	83	17	0	A
2O	99	1	0	A
2R	7	78	15	A
2S	91	9	0	A
2T	99	1	0	A
2U	94	6	0	A
2V	99	1	0	A
2J	82	18	0	A
2P	97	3	0	A
2Y	2	57	41	B
2Z	1	47	52	B
2AA	1	39	60	B
2AE	10	80	10	B
2AF	39	59	2	B
2AG	13	79	8	B
2AH	13	79	8	B
2AJ	1	49	49	B
2AK	6	77	18	B
2AM	29	68	3	B
2AO	15	78	7	B
2W	29	68	3	B
2X	18	76	5	B
2AB	9	79	12	B
2AC	4	73	22	B
2AP	29	68	3	B
2AQ	15	78	7	B
2AR	3	67	31	B
2AS	1	42	57	C
2AT	1	37	62	C
2AU	47	51	1	C
2AV	1	39	60	C
2AW	3	70	27	C
2AX	0	14	86	C
2AY	2	59	39	C
2AZ	0	4	96	C
2BA	0	0	100	C
2BB	0	1	99	C
2BC	0	1	99	C



9. 9th Place

Test Set

Confusion Matrix

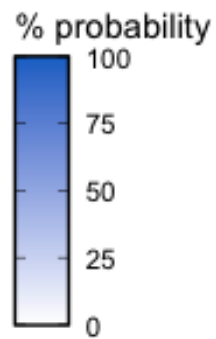


9. 9th Place

Test Set

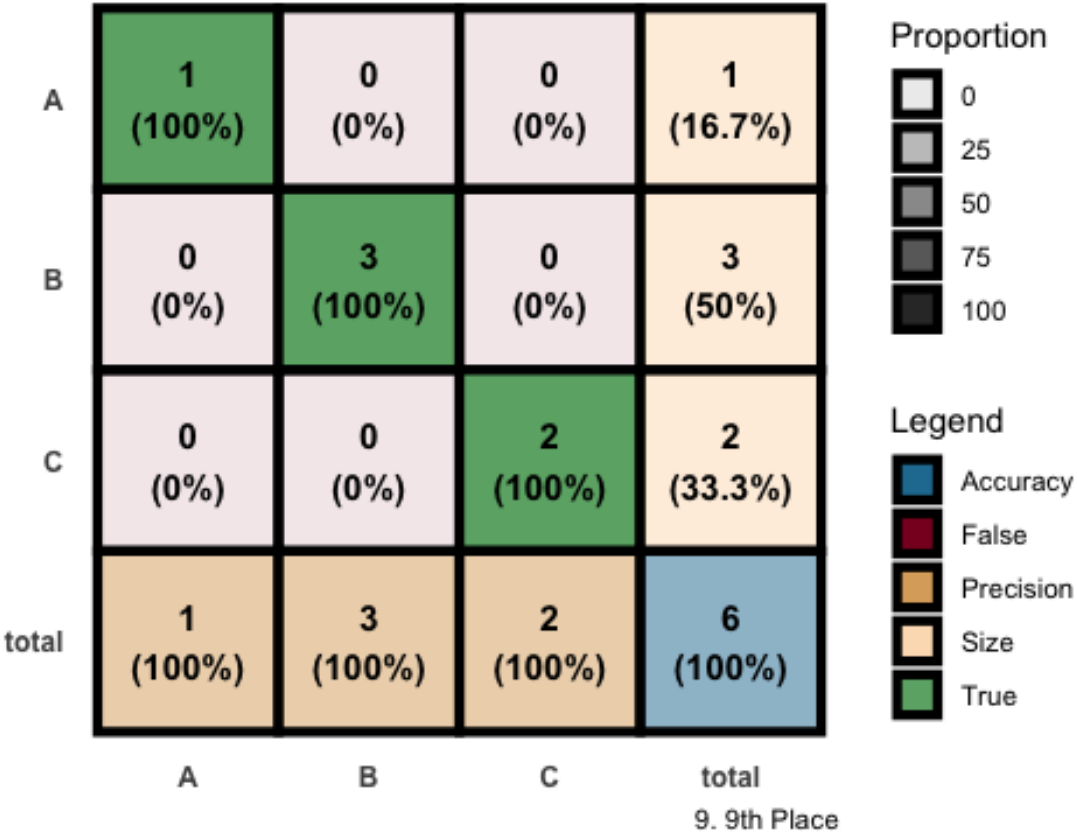
Probability Heatmap

	A	B	C	Exp
2B	77	23	0	A
2C	74	25	0	A
2D	99	1	0	A
2E	69	30	1	A
2F	73	26	0	A
2I	86	14	0	A
2K	100	0	0	A
2N	67	33	1	A
2Q	96	4	0	A
2AD	22	73	4	B
2AI	5	75	20	B
2AL	42	56	2	B
2AN	6	77	17	B



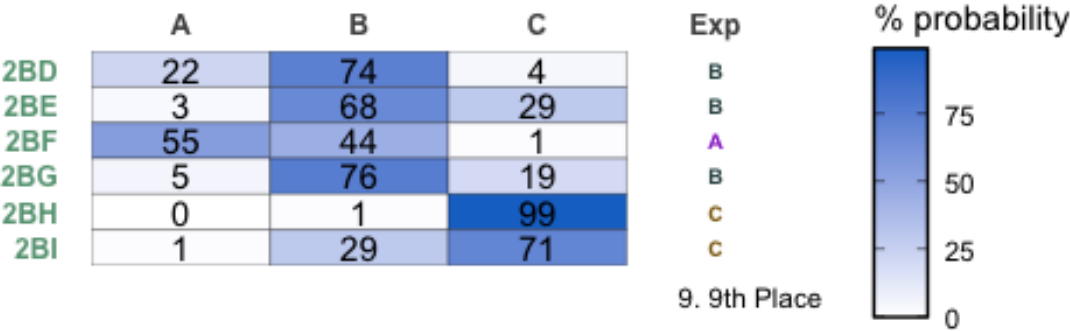
9. 9th Place

External Validation
Confusion Matrix



External Validation

Probability Heatmap



10th Place

Table:

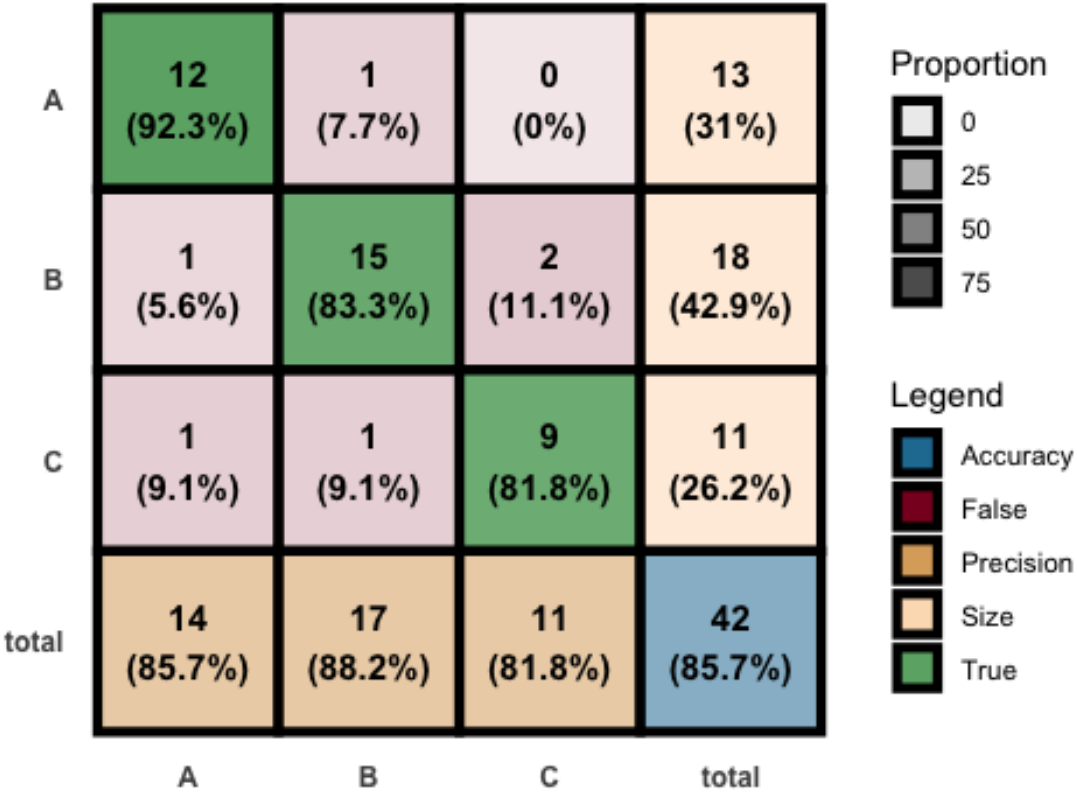
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.549

Model Coefficients

	x
-2-3-	-0.1368705
Total	0.9140378
NPA_1	32.9947500
NPA_3	-222.4549516

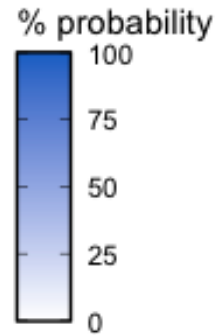
Training Set
Confusion Matrix



Training Set

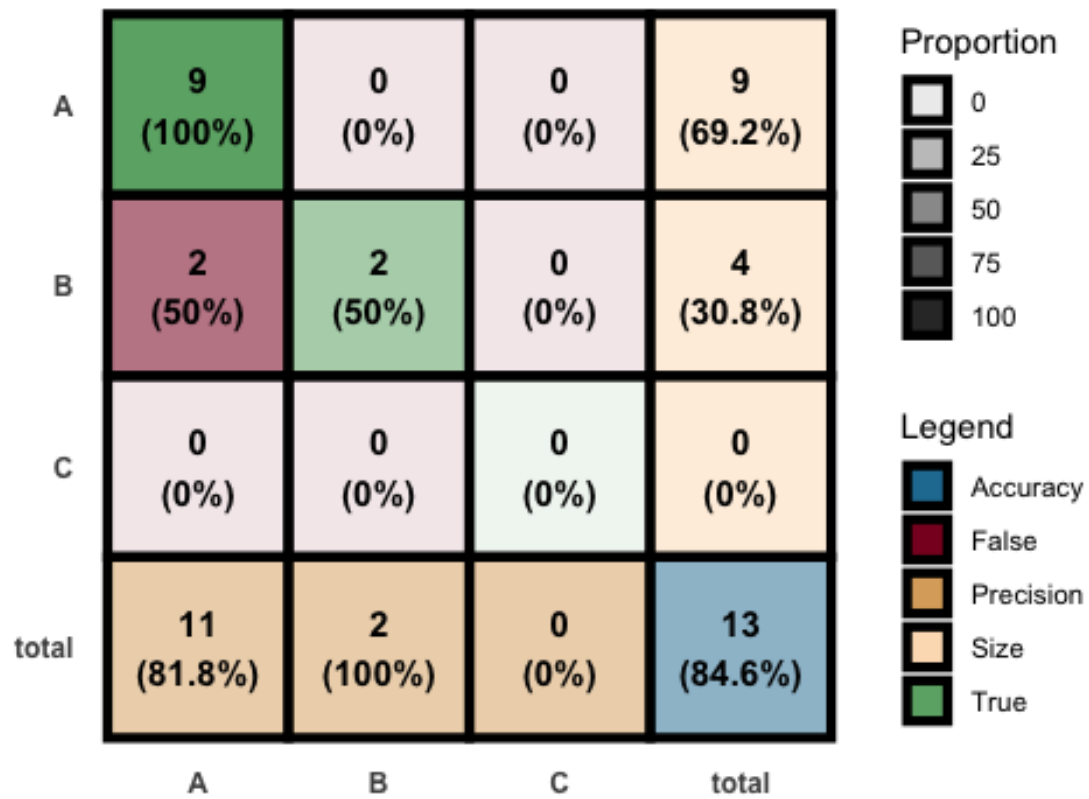
Probability Heatmap

	A	B	C	Exp
2A	68	32	0	A
2G	82	18	0	A
2H	97	3	0	A
2L	55	45	1	A
2M	76	23	0	A
2O	96	4	0	A
2R	20	76	4	A
2S	69	31	0	A
2T	98	2	0	A
2U	90	10	0	A
2V	96	4	0	A
2J	69	31	0	A
2P	98	2	0	A
2Y	7	81	12	B
2Z	1	54	45	B
2AA	0	29	71	B
2AE	11	82	7	B
2AF	23	74	3	B
2AG	20	76	4	B
2AH	22	75	3	B
2AJ	1	58	40	B
2AK	7	81	12	B
2AM	10	82	8	B
2AO	17	79	5	B
2W	57	42	1	B
2X	41	57	1	B
2AB	3	74	22	B
2AC	32	66	2	B
2AP	20	76	4	B
2AQ	11	82	7	B
2AR	1	47	52	B
2AS	1	42	57	C
2AT	0	21	79	C
2AU	69	31	0	C
2AV	0	18	82	C
2AW	3	73	24	C
2AX	0	8	92	C
2AY	1	48	52	C
2AZ	0	6	94	C
2BA	0	0	100	C
2BB	0	0	100	C
2BC	0	0	100	C



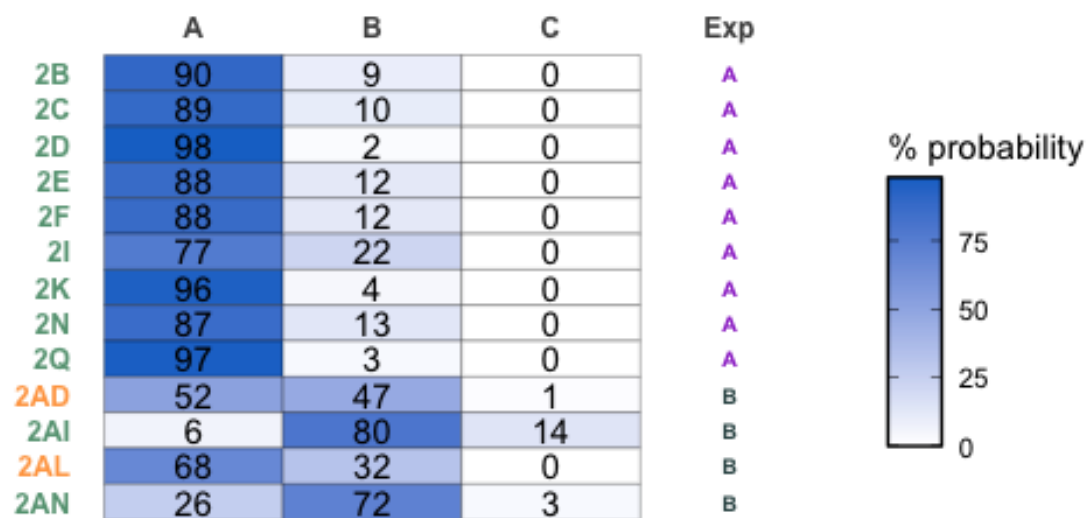
Test Set

Confusion Matrix



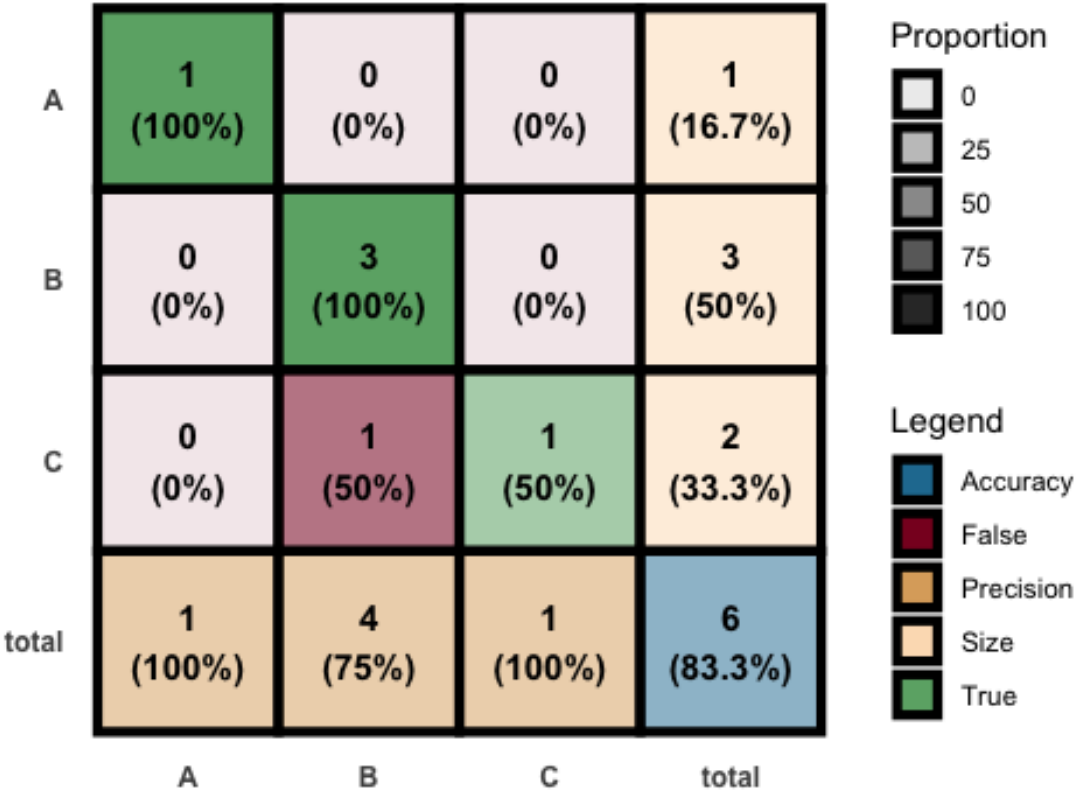
Test Set

Probability Heatmap



10. 10th Place

External Validation
Confusion Matrix



External Validation

Probability Heatmap

