

Pyridine Phosphination – 3 categories - Model Testing

Table of Contents

Model Search.....	1
Section Results Summary	1
1st Place.....	2
2nd Place.....	7
3rd Place.....	12
4th Place.....	17
5th Place.....	22
6th Place.....	27
7th Place.....	32
8th Place.....	37
9th Place.....	42
10th Place.....	47

Model Search

formula	McFadden R2
class ~ para_P + NPA_6_SM + Avg_NPA_SM	0.703
class ~ NPA_6_SM + para_SM + fr_C_O	0.662
class ~ NPA_6_SM + para_SM + fr_C_O_noCOO	0.662
class ~ Dist_(P,C)_P + para_SM + Iso.Polar_SM	0.651
class ~ para_P + NPA_6_SM + NPA_sum_SM	0.649
class ~ Dist_(P,C)_P + NPA_6_SM + para_SM	0.649
class ~ NPA_6_SM + Avg_NPA_SM + para_SM	0.646
class ~ NPA_1_P + Dist_(P,C)_P + para_SM	0.627
class ~ NPA_1_P + para_SM + fr_pyridine	0.626
class ~ NPA_6_SM + NPA_sum_SM + para_SM	0.623

Section Results Summary

	Training Accuracy	Test Accuracy
1. 1st Place	91.43	54.5
2. 2nd Place	85.71	27.3
3. 3rd Place	85.71	27.3
4. 4th Place	85.71	72.7
5. 5th Place	88.57	63.6
6. 6th Place	88.57	81.8
7. 7th Place	82.86	54.5
8. 8th Place	88.57	63.6

	Training Accuracy	Test Accuracy
9. 9th Place	82.86	45.5
10. 10th Place	85.71	54.5

1st Place

Table:

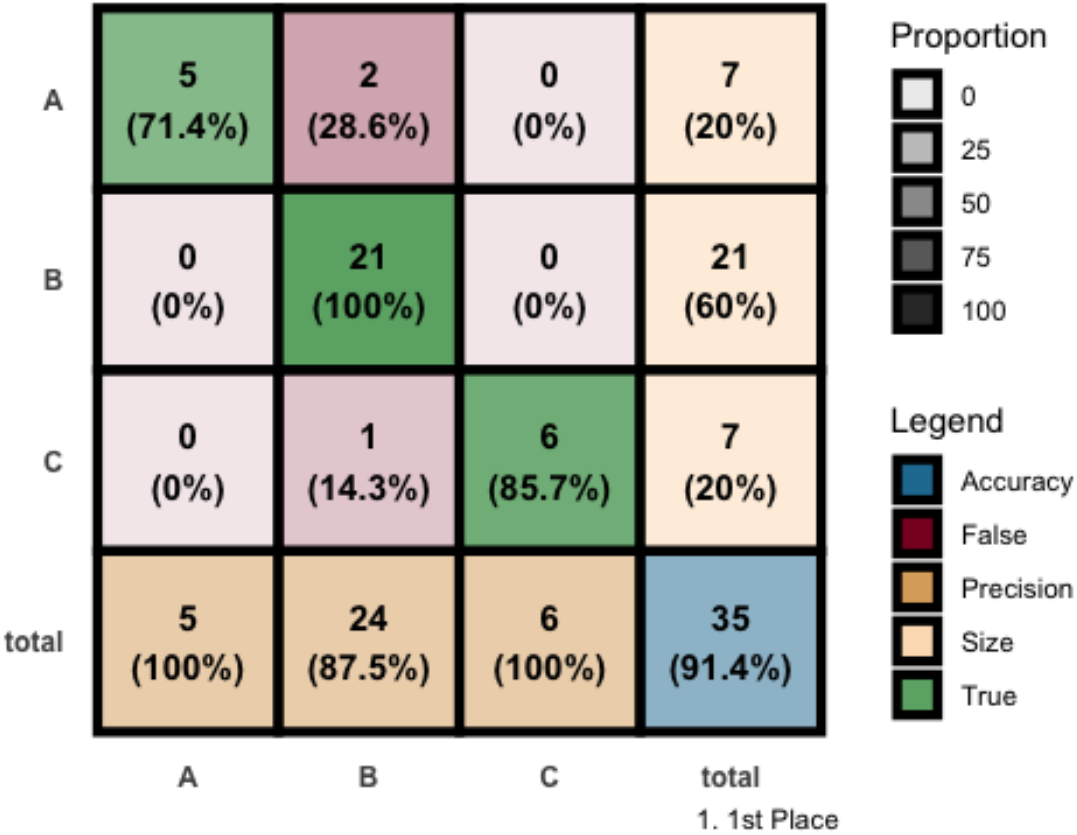
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
91.43%	0.703

Model Coefficients

	x
para_P	0.1124591
NPA_6_SM	35.1047354
Avg_NPA_SM	-98.8554566

Training Set
Confusion Matrix

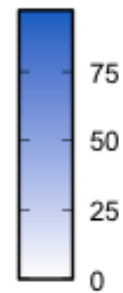


Training Set

Probability Heatmap

	A	B	C	Exp
A	17	83	0	A
AF	83	17	0	A
B	66	34	0	A
C	96	4	0	A
D	95	5	0	A
E	38	62	0	A
F	95	5	0	A
AA	0	75	25	B
AB	13	86	1	B
AD	25	74	0	B
AH	1	87	12	B
AJ	1	93	6	B
AK	0	73	27	B
AL	4	94	2	B
AM	0	82	18	B
H	40	60	0	B
I	16	83	0	B
J	2	94	4	B
L	2	94	4	B
M	4	94	2	B
O	7	92	1	B
P	19	81	0	B
Q	40	60	0	B
T	38	62	0	B
U	18	81	0	B
V	15	85	0	B
Y	1	93	6	B
Z	5	94	2	B
AN	0	40	60	C
AP	0	5	95	C
AQ	0	60	40	C
AR	0	5	95	C
AS	0	3	97	C
AT	0	6	94	C
AU	0	42	58	C

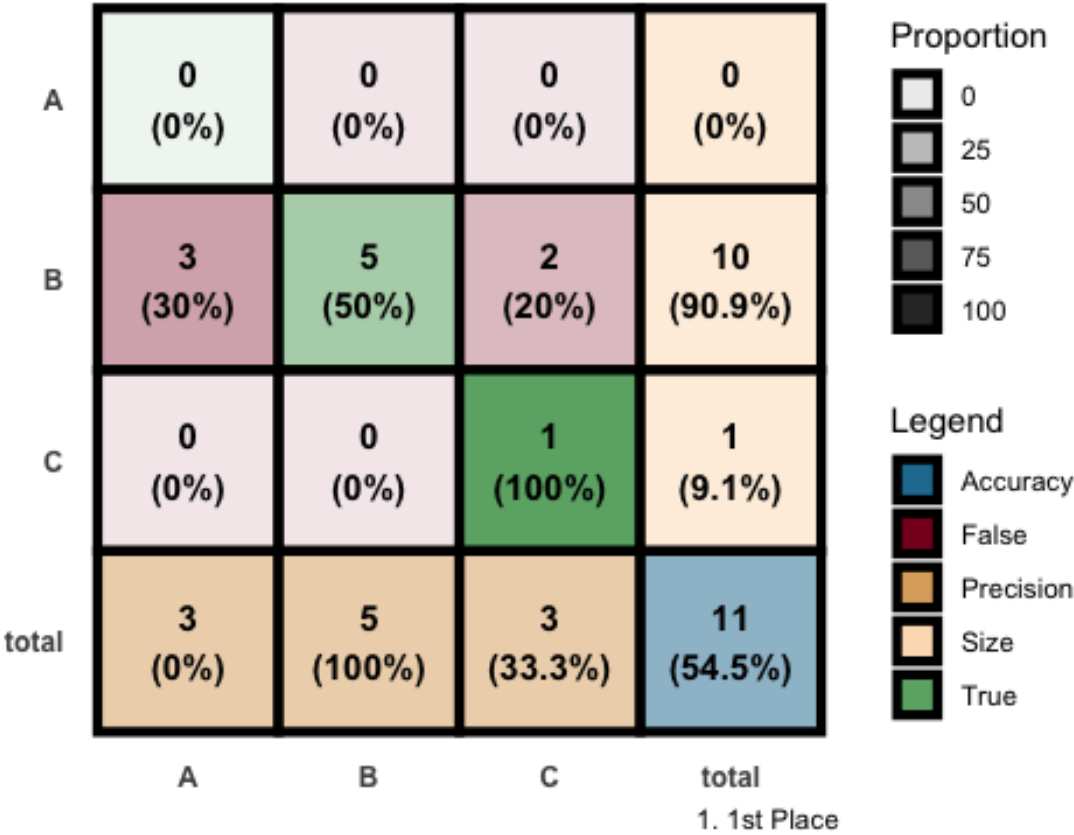
% probability



1. 1st Place

Test Set

Confusion Matrix



Test Set

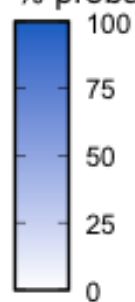
Probability Heatmap

	A	B	C
AC	14	85	1
AE	12	88	1
AI	73	27	0
G	2	94	4
K	0	42	58
N	100	0	0
R	40	60	0
S	86	14	0
W	0	26	74
X	25	75	0
AO	0	11	89

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



1. 1st Place

2nd Place

Table:

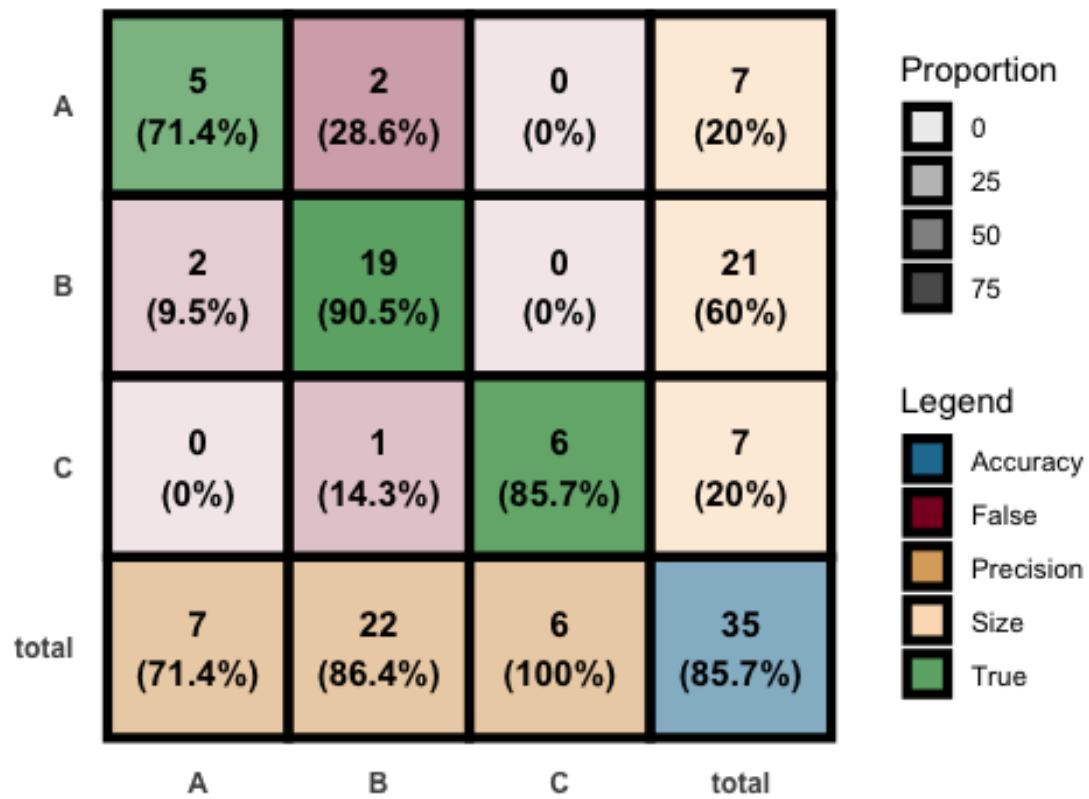
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.662

Model Coefficients

	x
NPA_6_SM	39.6636070
para_SM	0.1810645
fr_C_O1	4.4955505
fr_C_O2	17.9743375

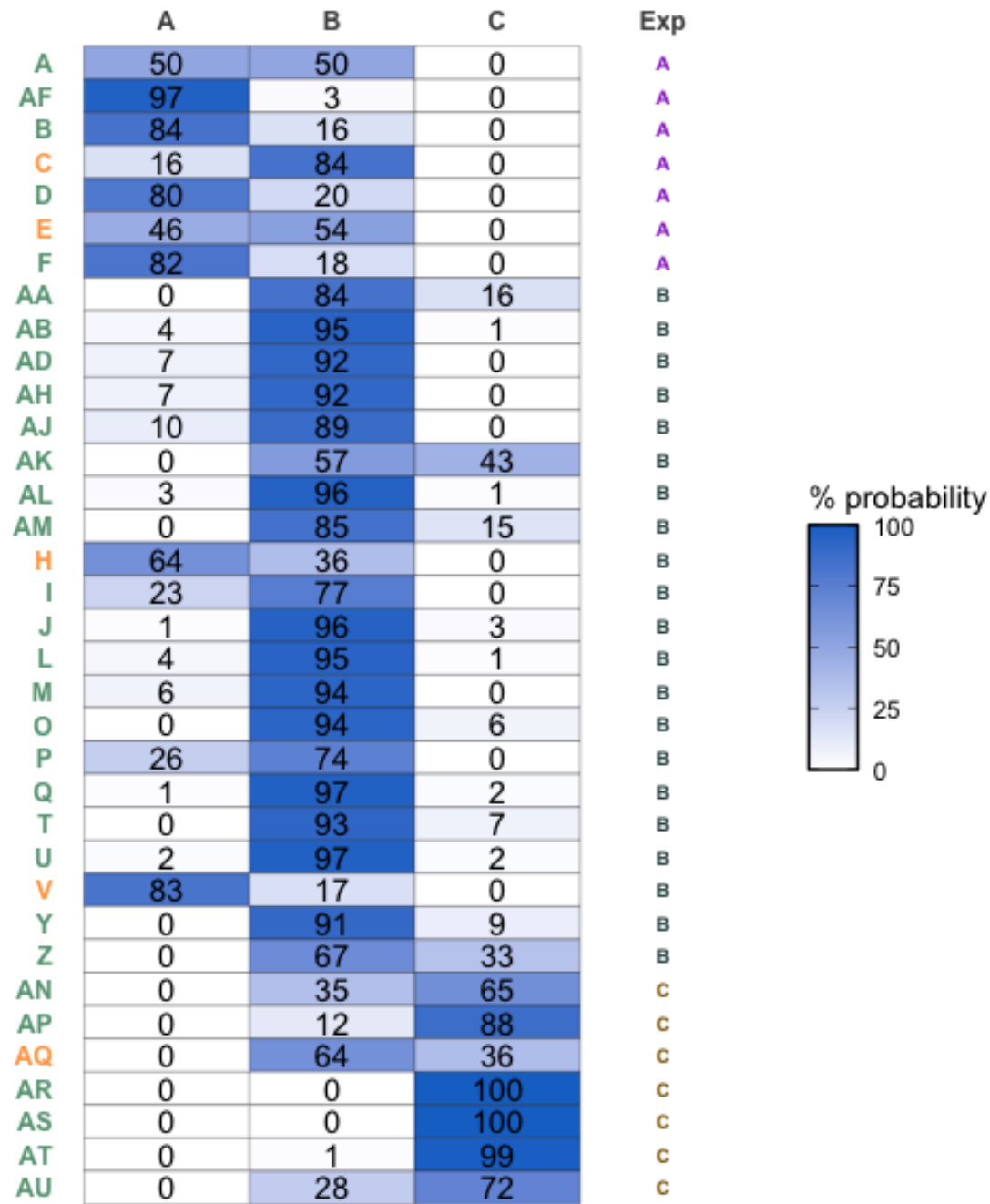
Training Set
Confusion Matrix



2. 2nd Place

Training Set

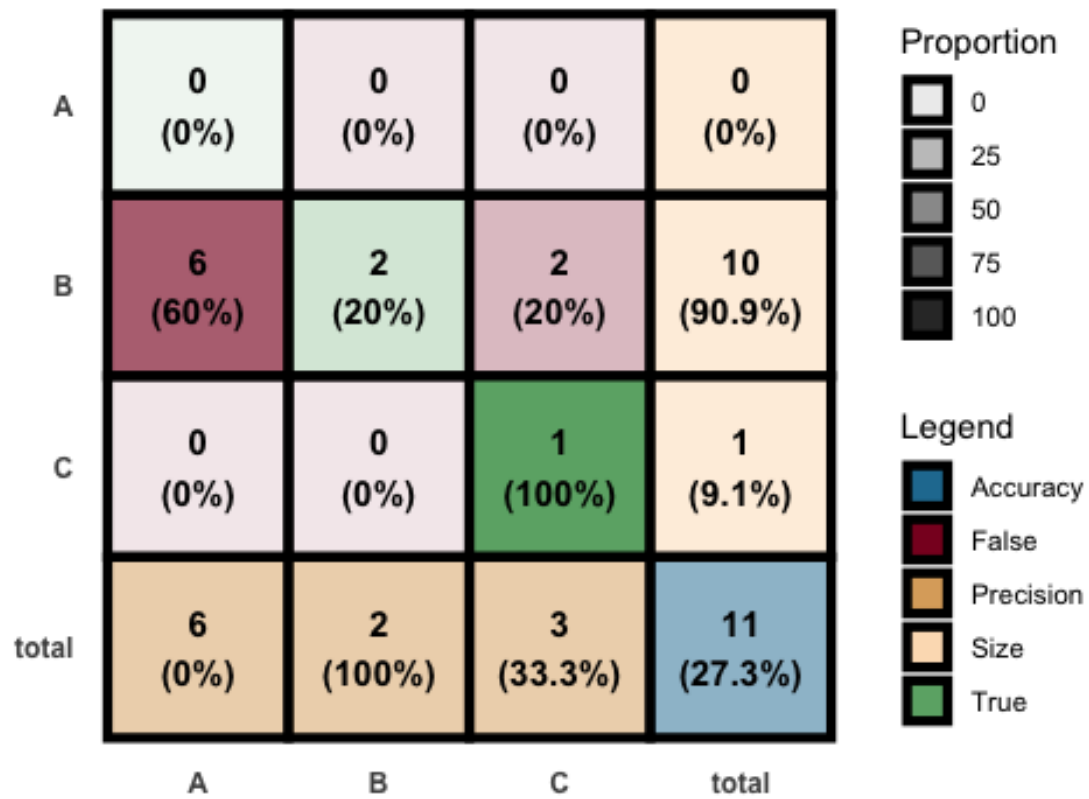
Probability Heatmap



2. 2nd Place

Test Set

Confusion Matrix



Test Set

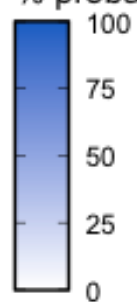
Probability Heatmap

	A	B	C
AC	81	19	0
AE	0	35	65
AI	51	49	0
G	1	96	3
K	0	61	39
N	100	0	0
R	64	36	0
S	65	35	0
W	0	0	100
X	75	25	0
AO	0	34	66

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



2. 2nd Place

3rd Place

Table:

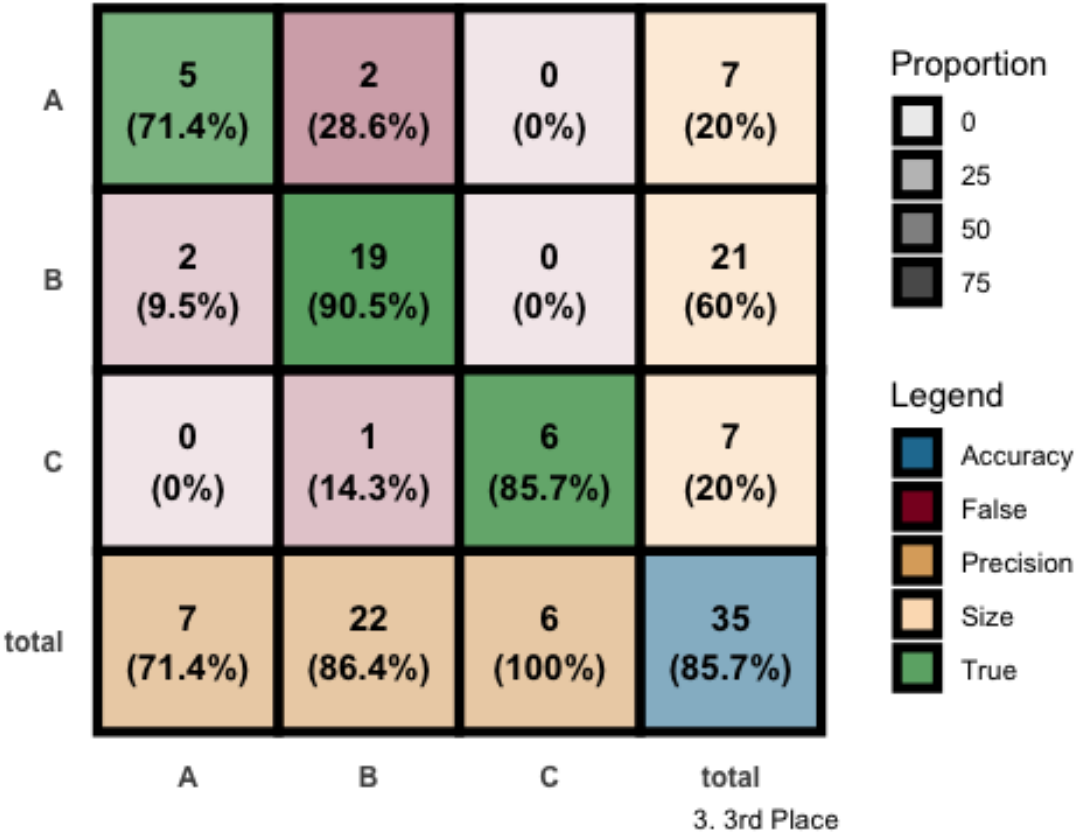
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.662

Model Coefficients

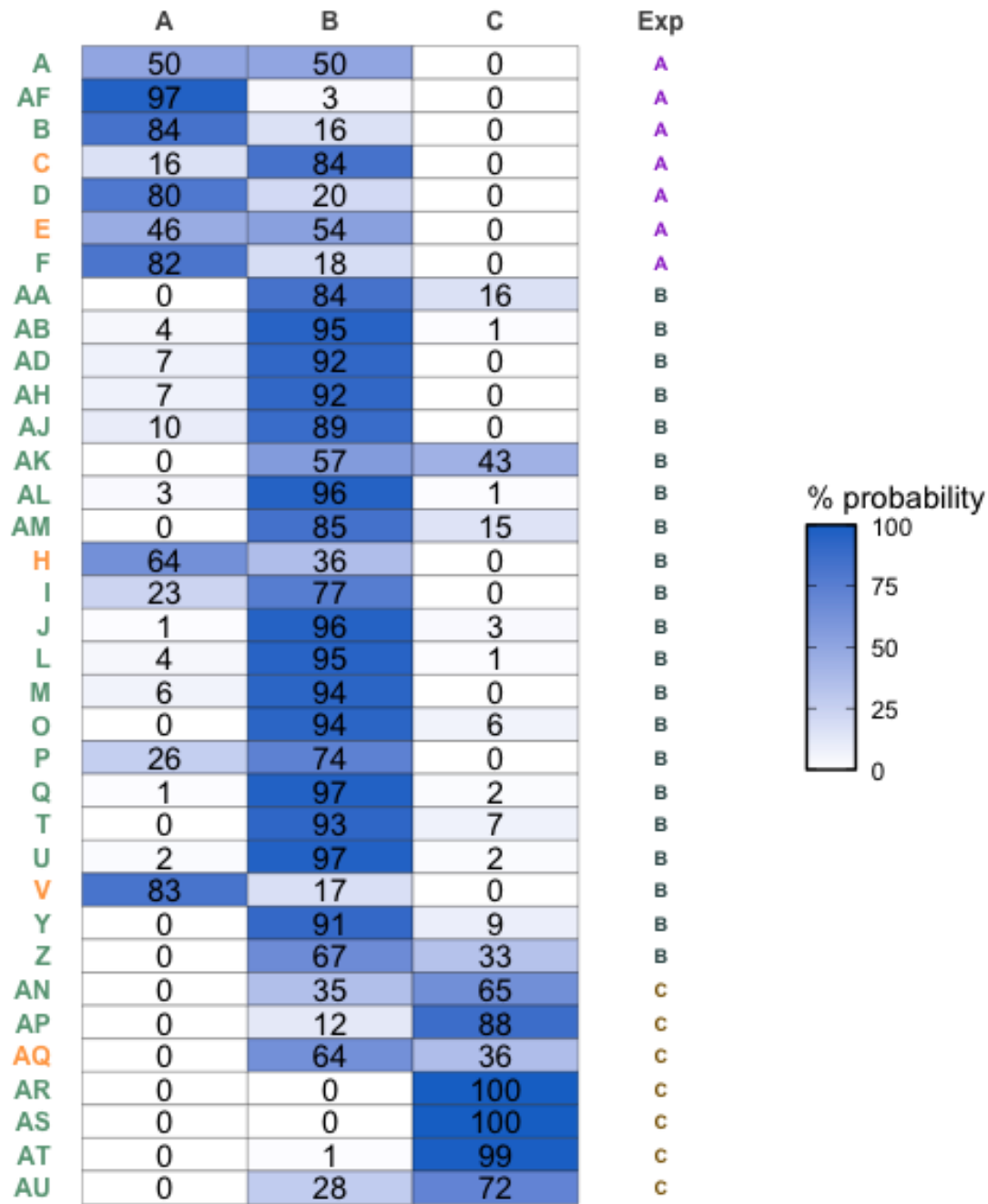
	x
NPA_6_SM	39.6636070
para_SM	0.1810645
fr_C_O_noC001	4.4955505
fr_C_O_noC002	17.9743375

Training Set
Confusion Matrix



Training Set

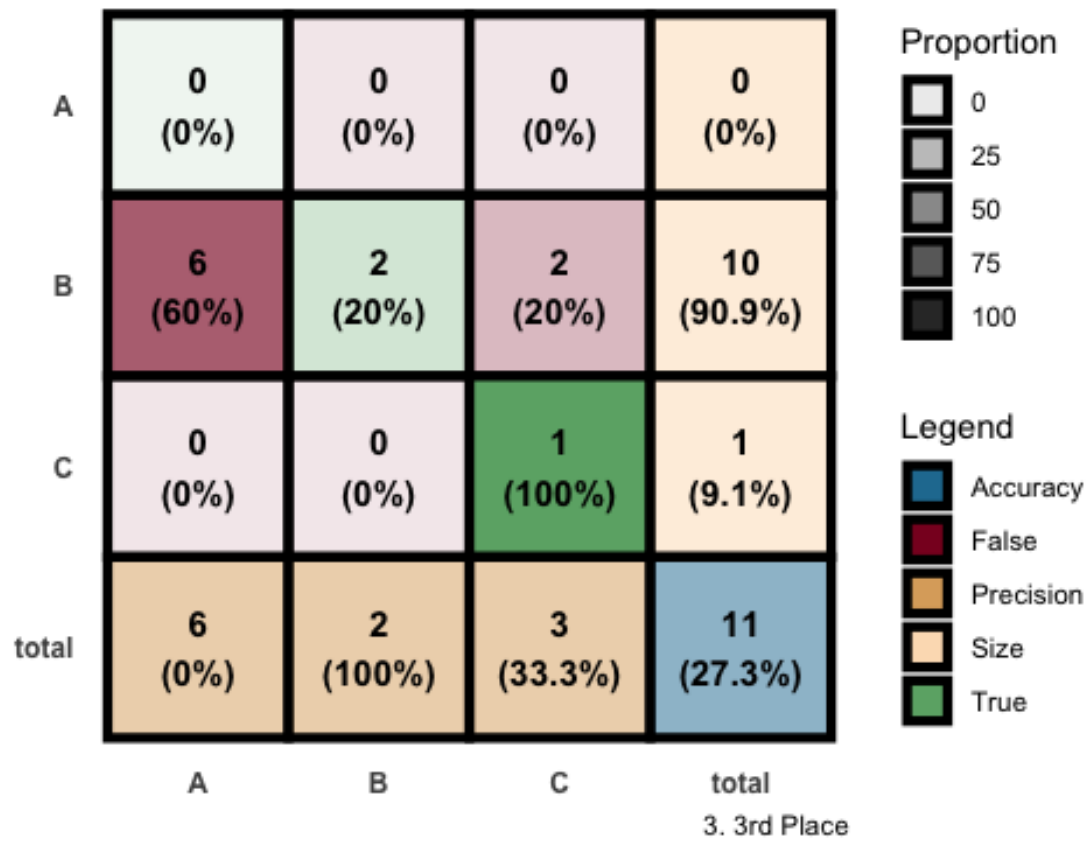
Probability Heatmap



3. 3rd Place

Test Set

Confusion Matrix



Test Set

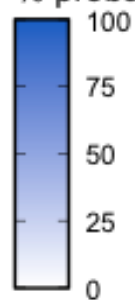
Probability Heatmap

	A	B	C
AC	81	19	0
AE	0	35	65
AI	51	49	0
G	1	96	3
K	0	61	39
N	100	0	0
R	64	36	0
S	65	35	0
W	0	0	100
X	75	25	0
AO	0	34	66

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



3. 3rd Place

4th Place

Table:

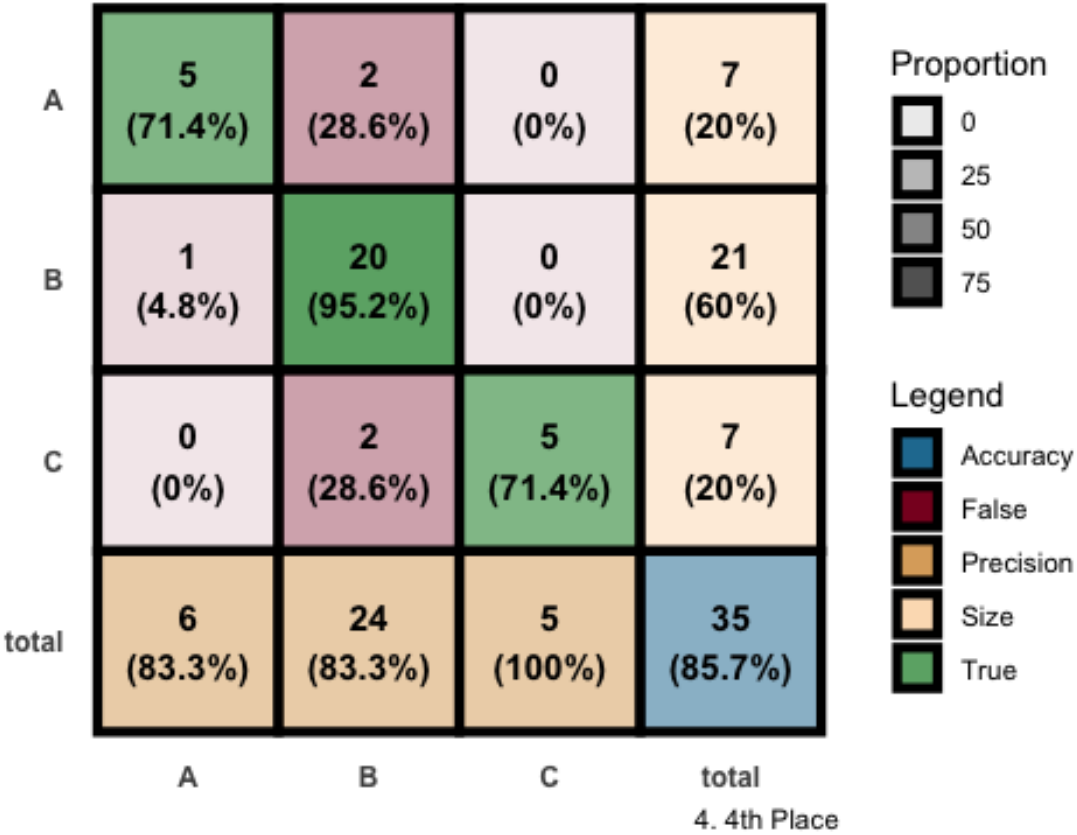
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.651

Model Coefficients

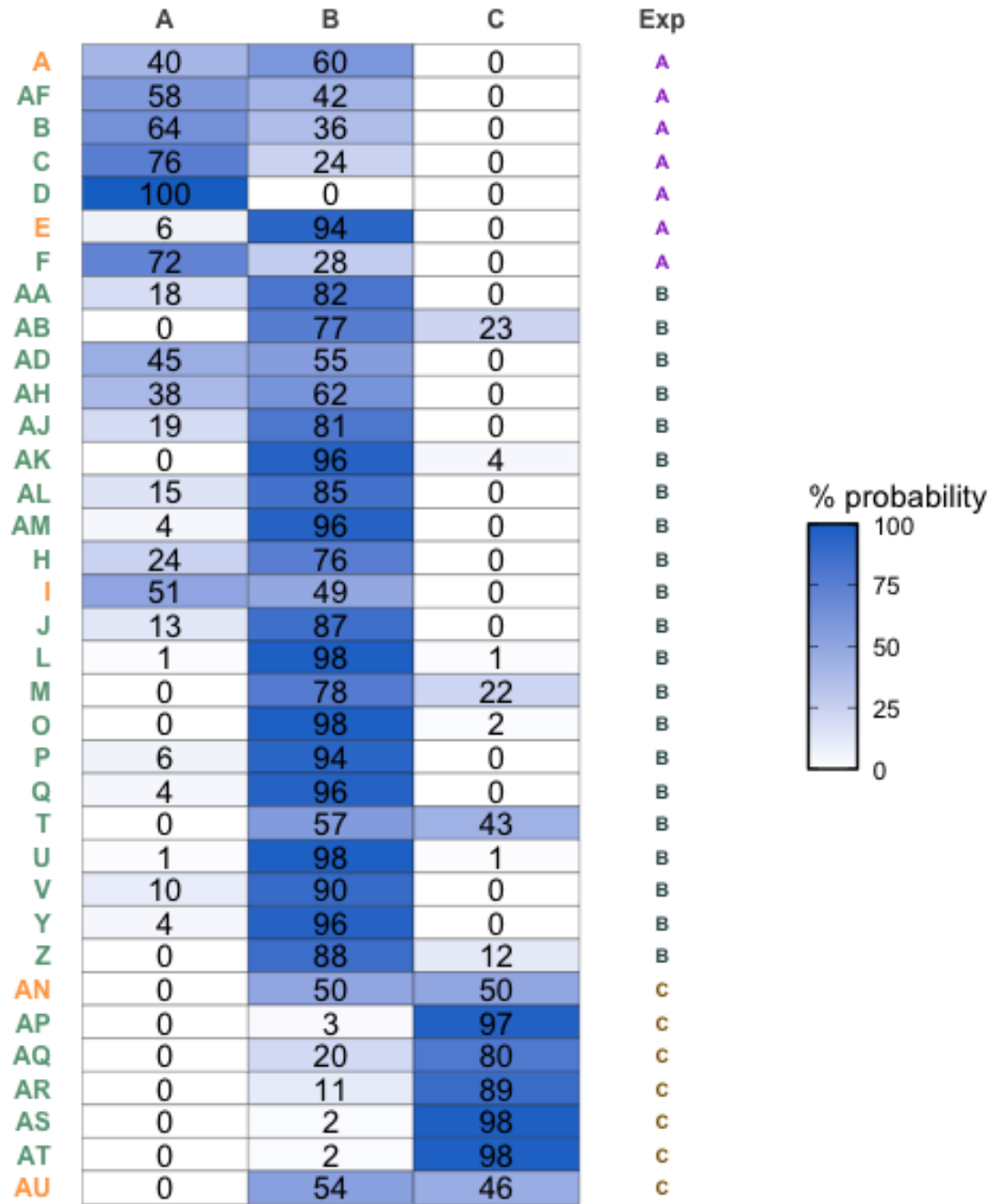
	x
Dist_(P,C)_P	-257.2488454
para_SM	0.1905674
Iso.Polar_SM	0.0329815

Training Set
Confusion Matrix



Training Set

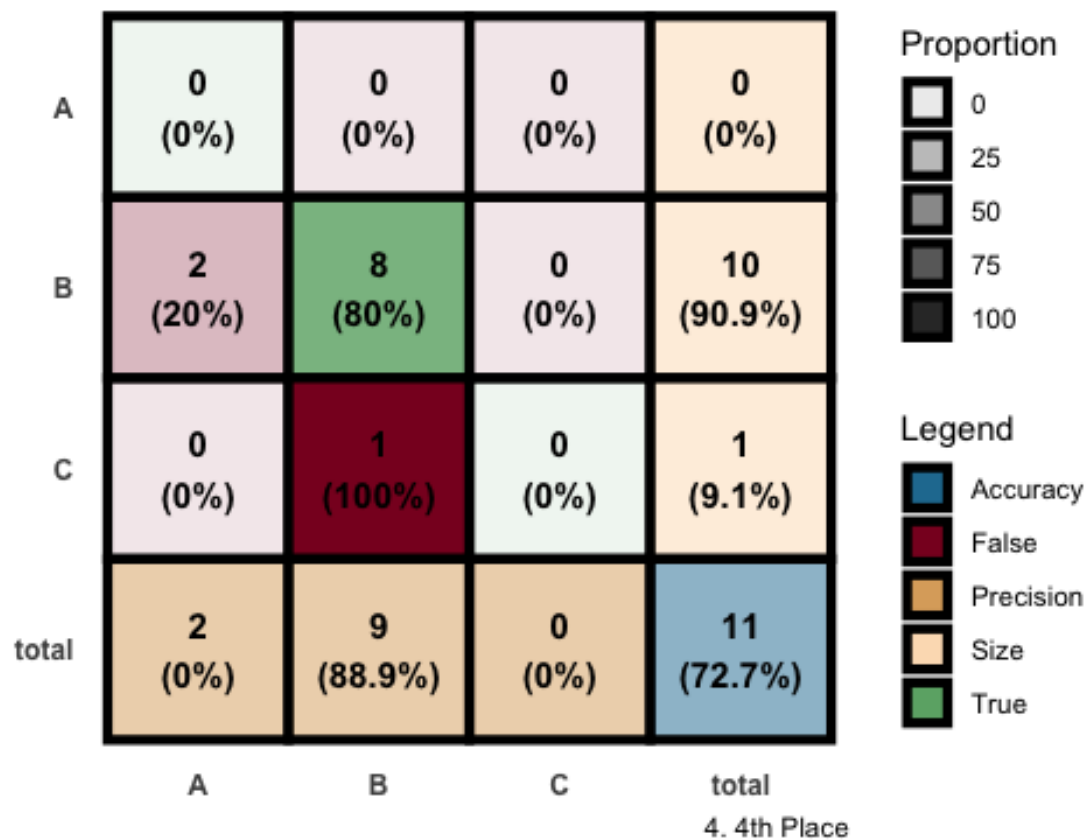
Probability Heatmap



4. 4th Place

Test Set

Confusion Matrix



Test Set

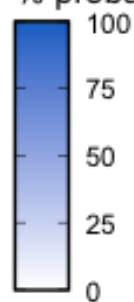
Probability Heatmap

	A	B	C
AC	1	98	1
AE	0	90	10
AI	0	73	27
G	8	92	0
K	12	88	0
N	98	2	0
R	9	91	0
S	100	0	0
W	1	98	1
X	0	97	3
AO	0	69	31

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



4. 4th Place

5th Place

Table:

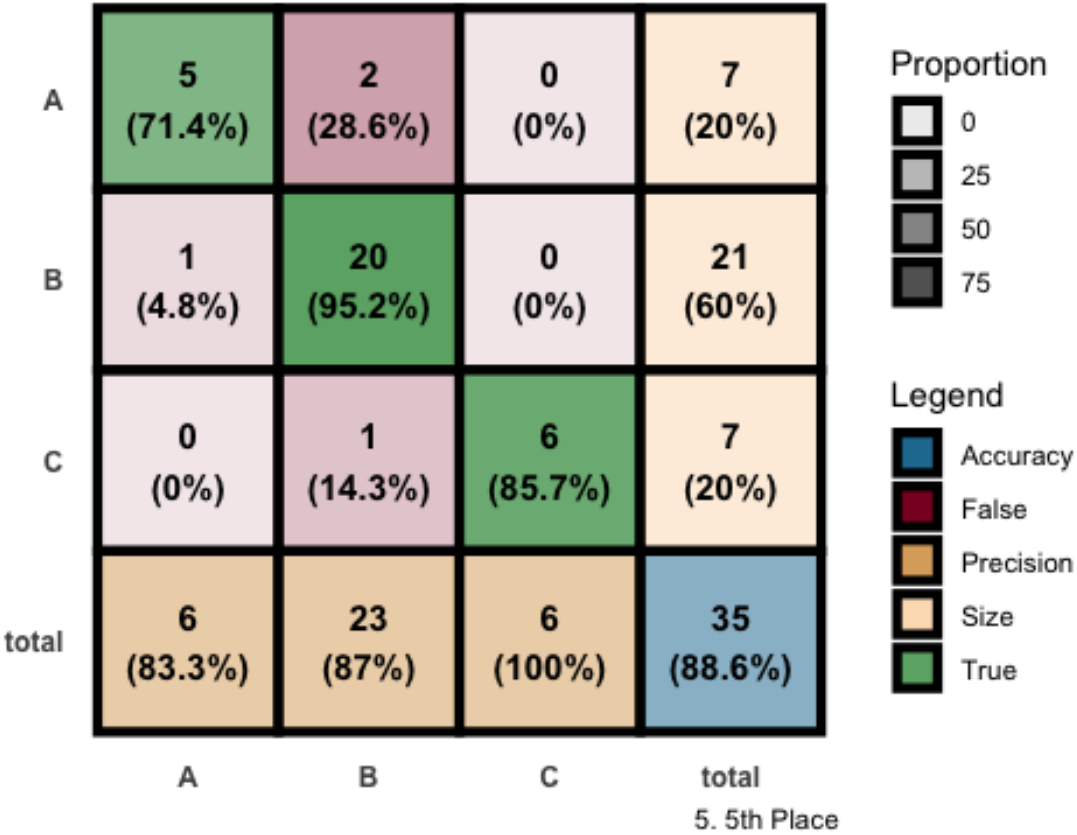
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
88.57%	0.649

Model Coefficients

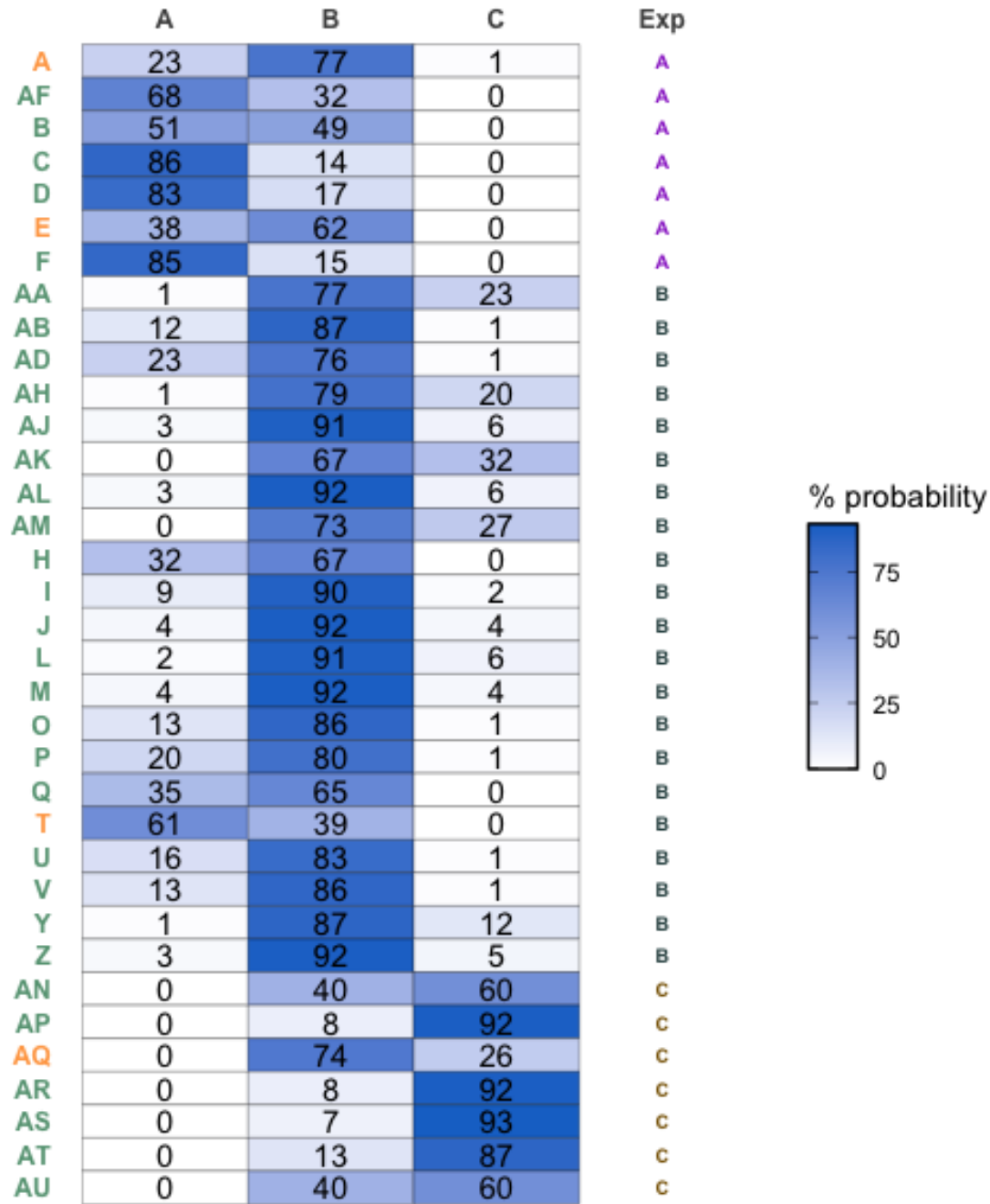
	x
para_P	0.0774822
NPA_6_SM	31.0525158
NPA_sum_SM	-13.5126258

Training Set
Confusion Matrix



Training Set

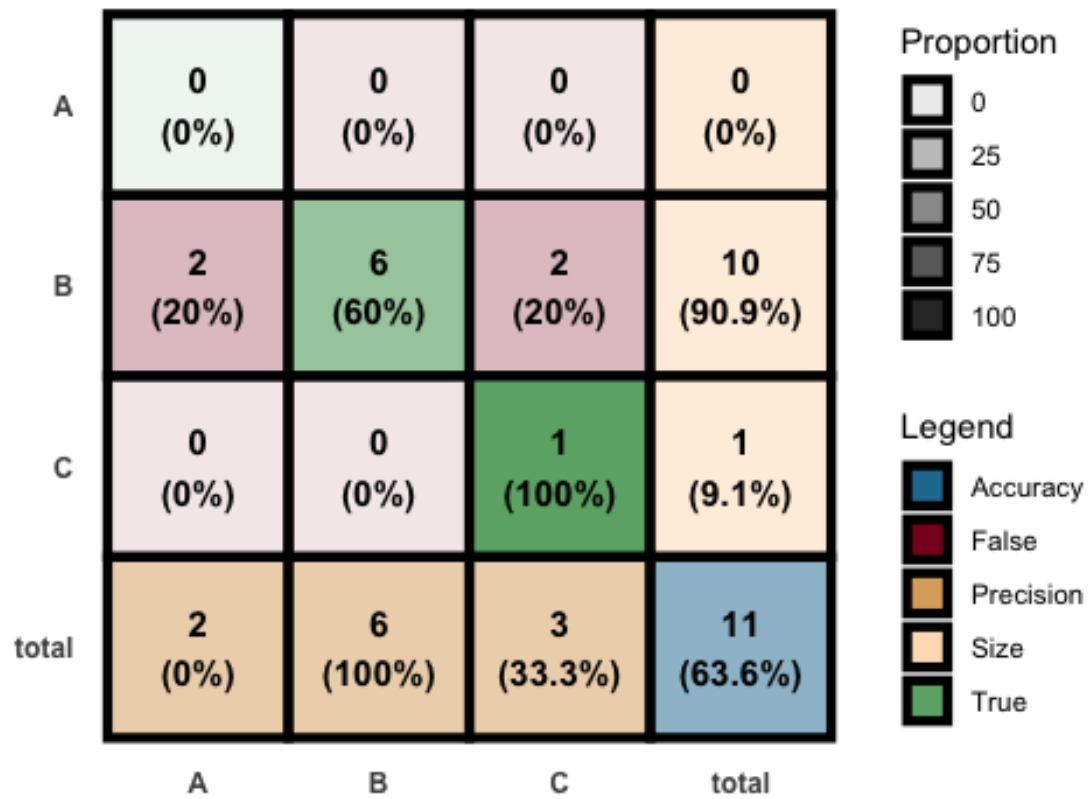
Probability Heatmap



5. 5th Place

Test Set

Confusion Matrix



5. 5th Place

Test Set

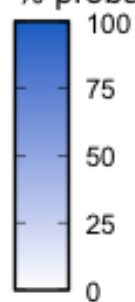
Probability Heatmap

	A	B	C
AC	10	88	1
AE	14	85	1
AI	68	32	0
G	4	92	4
K	0	39	61
N	100	0	0
R	27	73	0
S	43	56	0
W	0	23	77
X	21	78	1
AO	0	17	82

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



5. 5th Place

6th Place

Table:

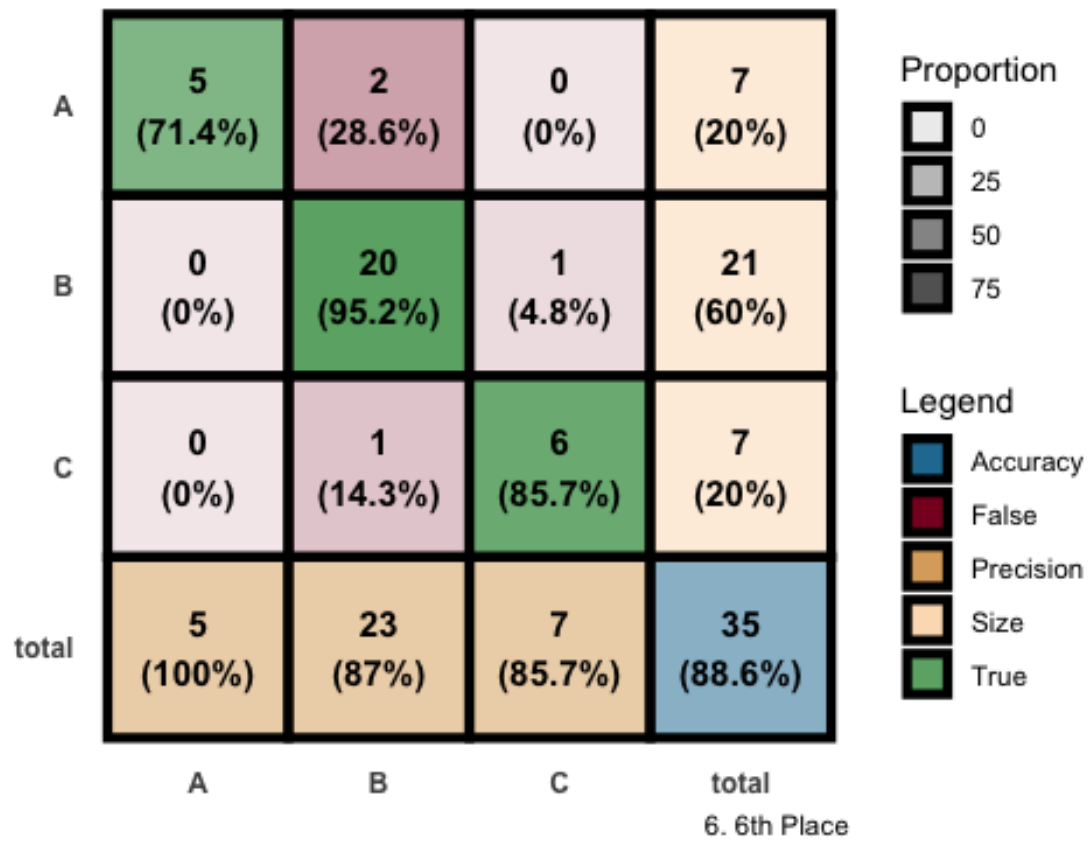
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
88.57%	0.649

Model Coefficients

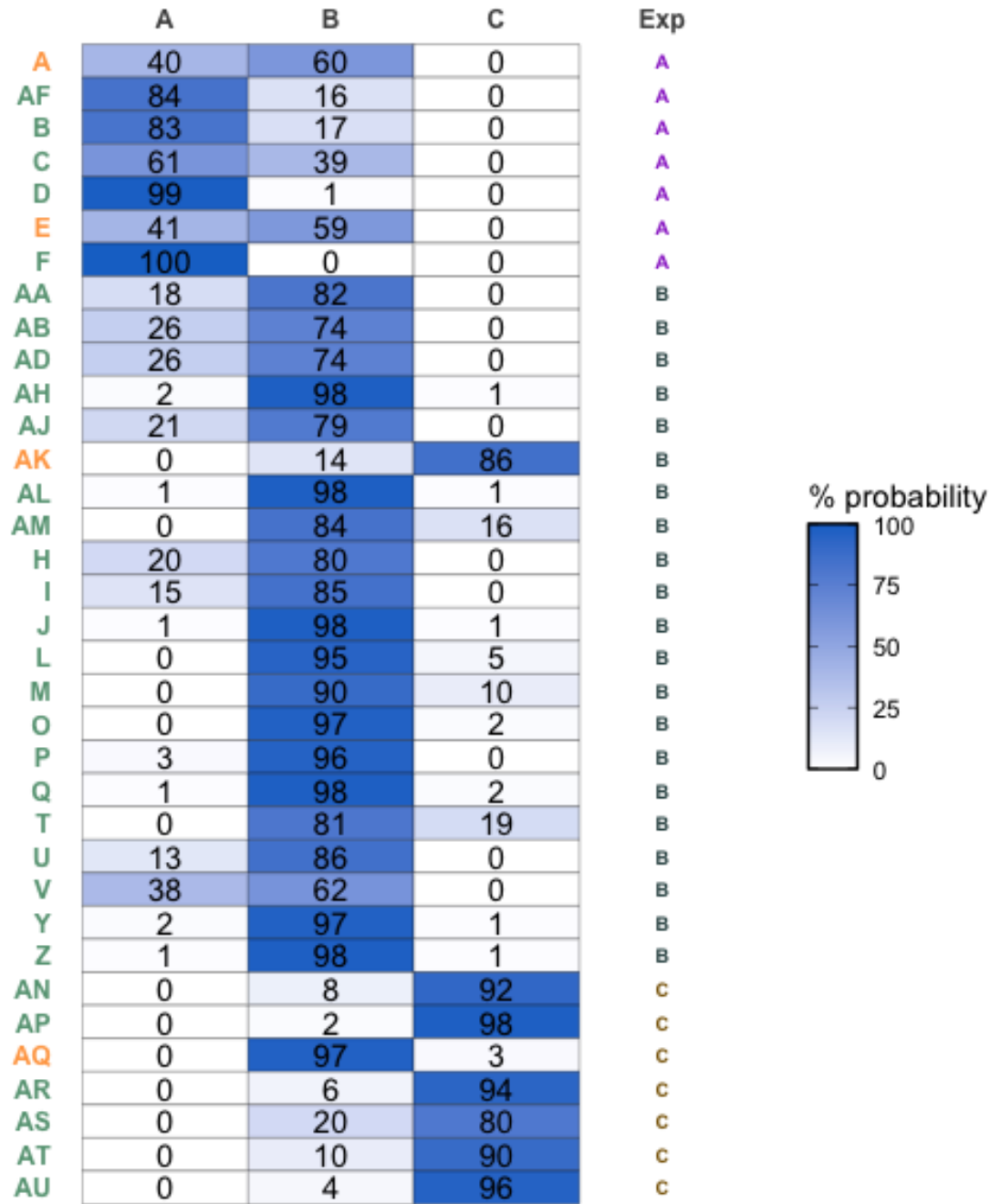
	x
Dist_(P,C)_P	-181.4237850
NPA_6_SM	33.1536447
para_SM	0.2071308

Training Set Confusion Matrix



Training Set

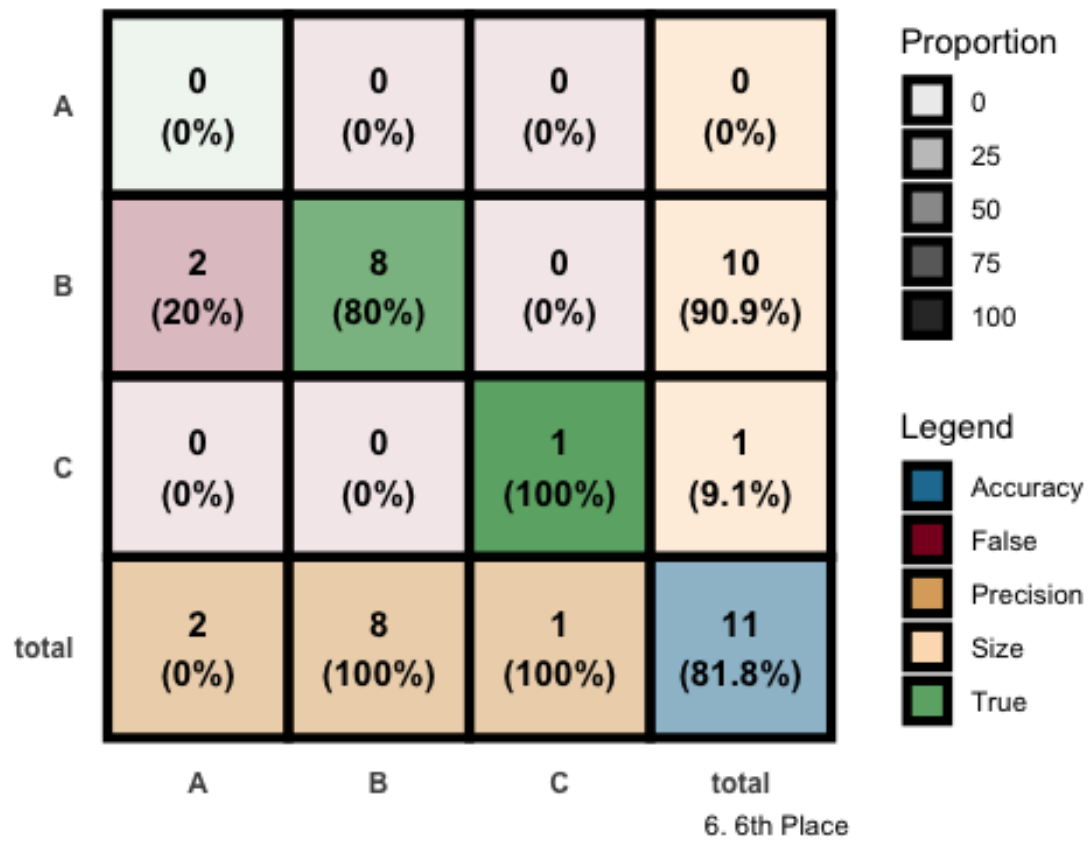
Probability Heatmap



6. 6th Place

Test Set

Confusion Matrix



Test Set

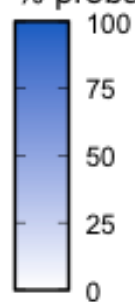
Probability Heatmap

	A	B	C
AC	12	88	0
AE	0	96	4
AI	1	98	2
G	38	62	0
K	0	83	17
N	100	0	0
R	23	77	0
S	80	20	0
W	0	88	12
X	24	76	0
AO	0	6	94

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



6. 6th Place

7th Place

Table:

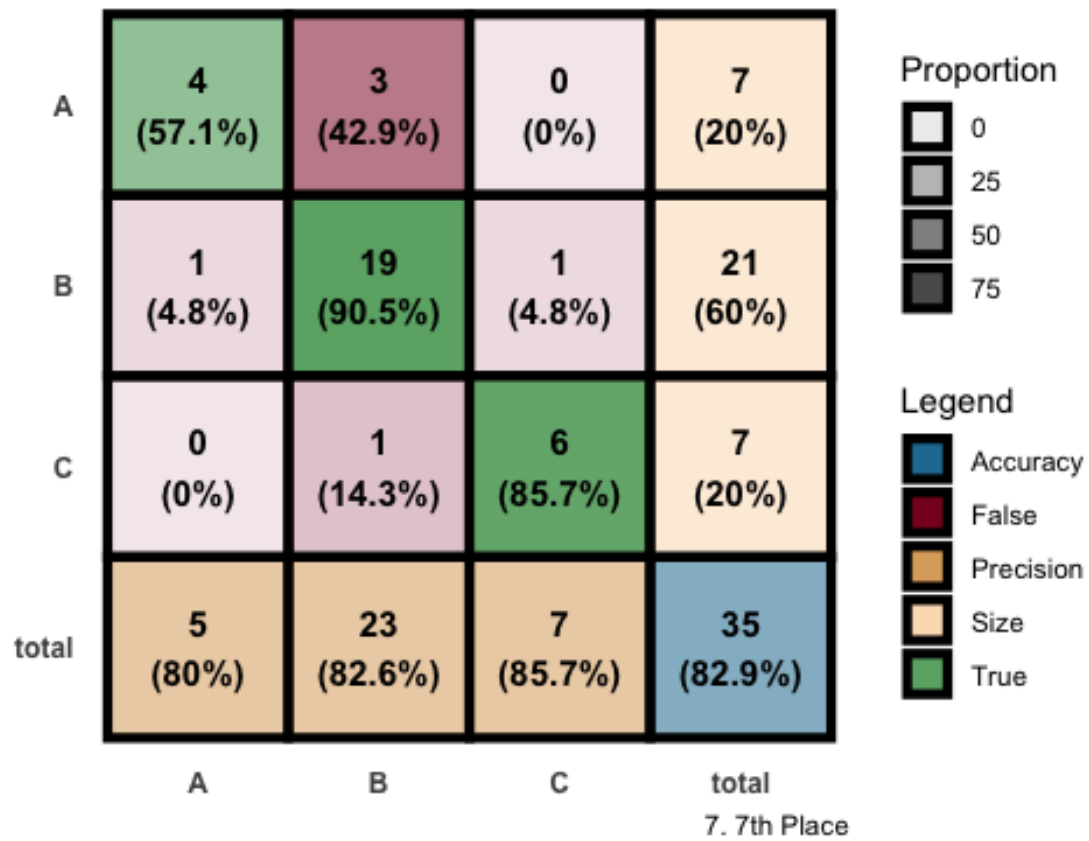
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
82.86%	0.646

Model Coefficients

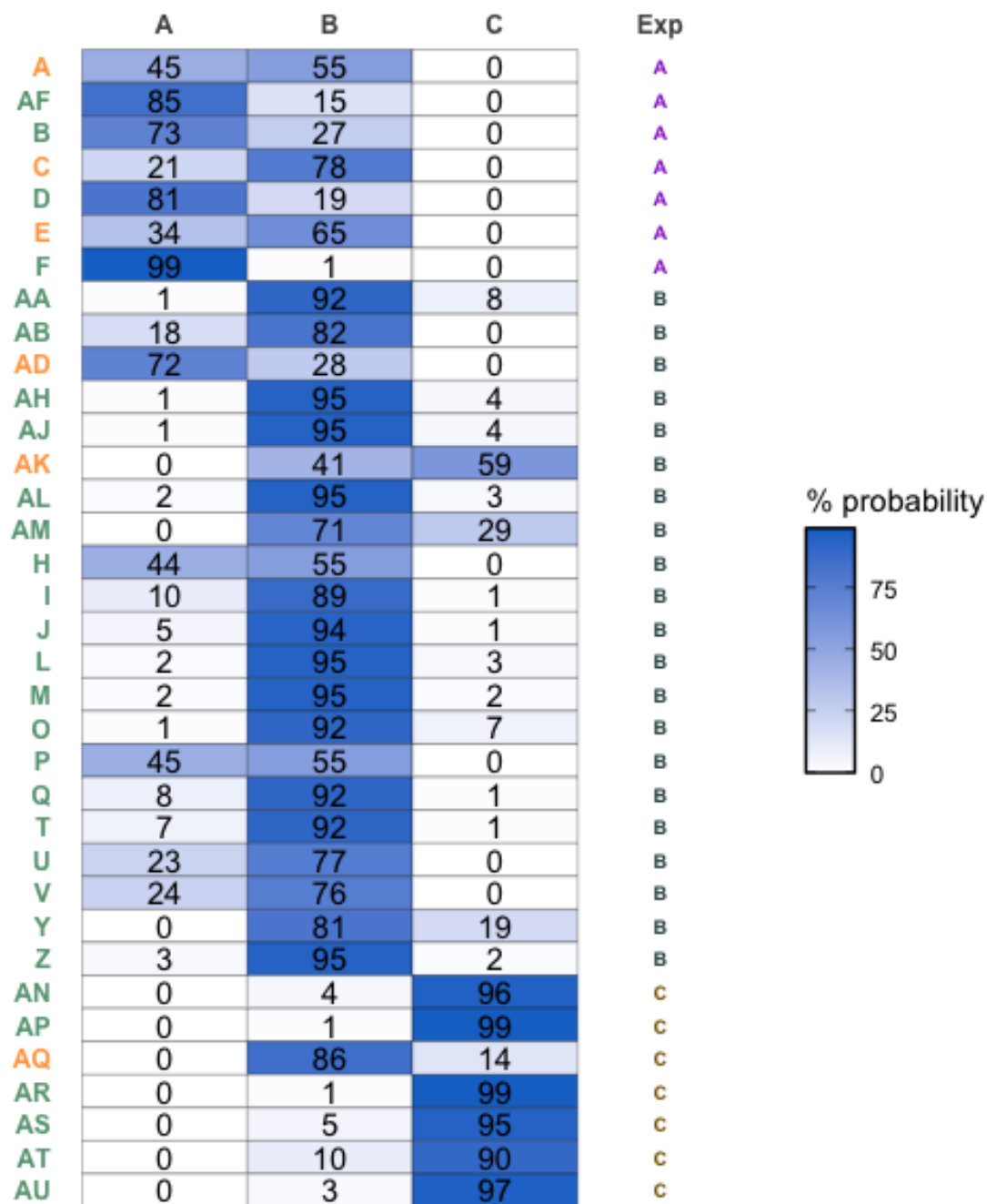
	x
NPA_6_SM	37.2573966
Avg_NPA_SM	-69.5656894
para_SM	0.1550816

Training Set Confusion Matrix



Training Set

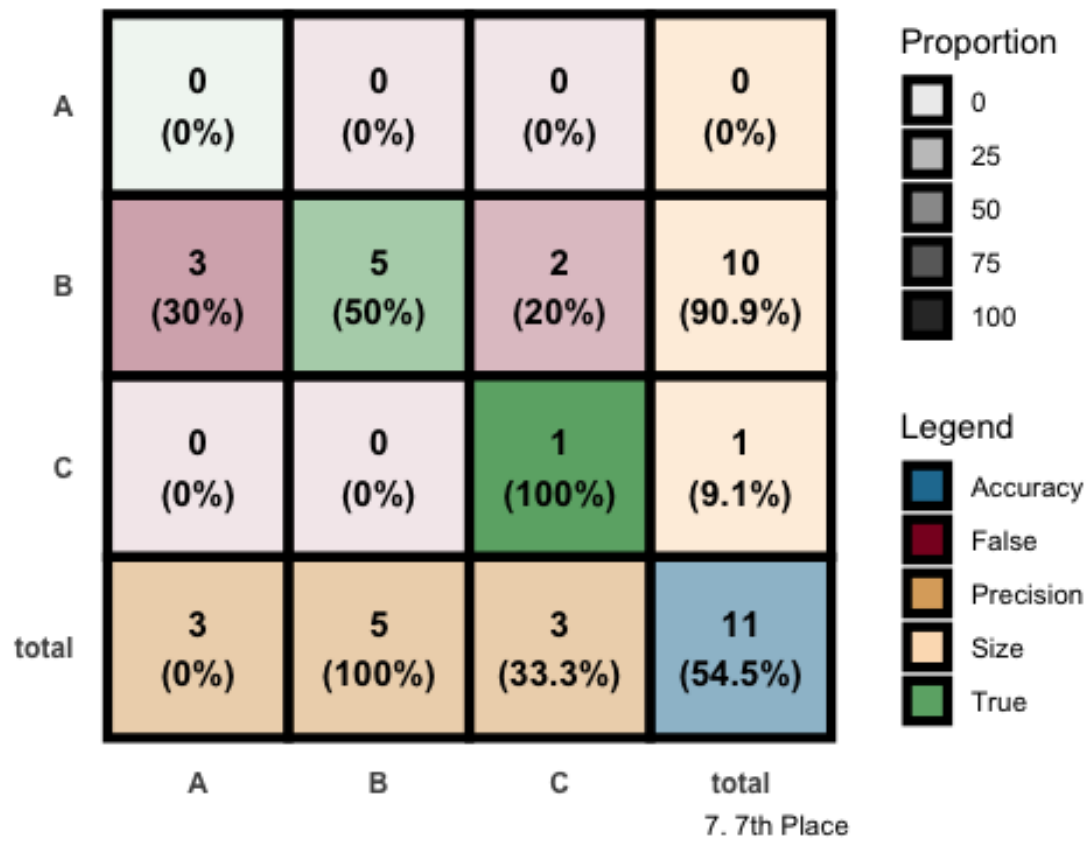
Probability Heatmap



7. 7th Place

Test Set

Confusion Matrix



Test Set

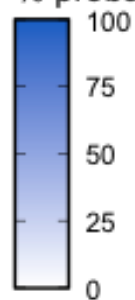
Probability Heatmap

	A	B	C
AC	17	82	0
AE	8	91	1
AI	63	37	0
G	6	93	1
K	0	29	71
N	100	0	0
R	11	88	0
S	83	17	0
W	0	24	76
X	20	80	0
AO	0	3	97

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



7. 7th Place

8th Place

Table:

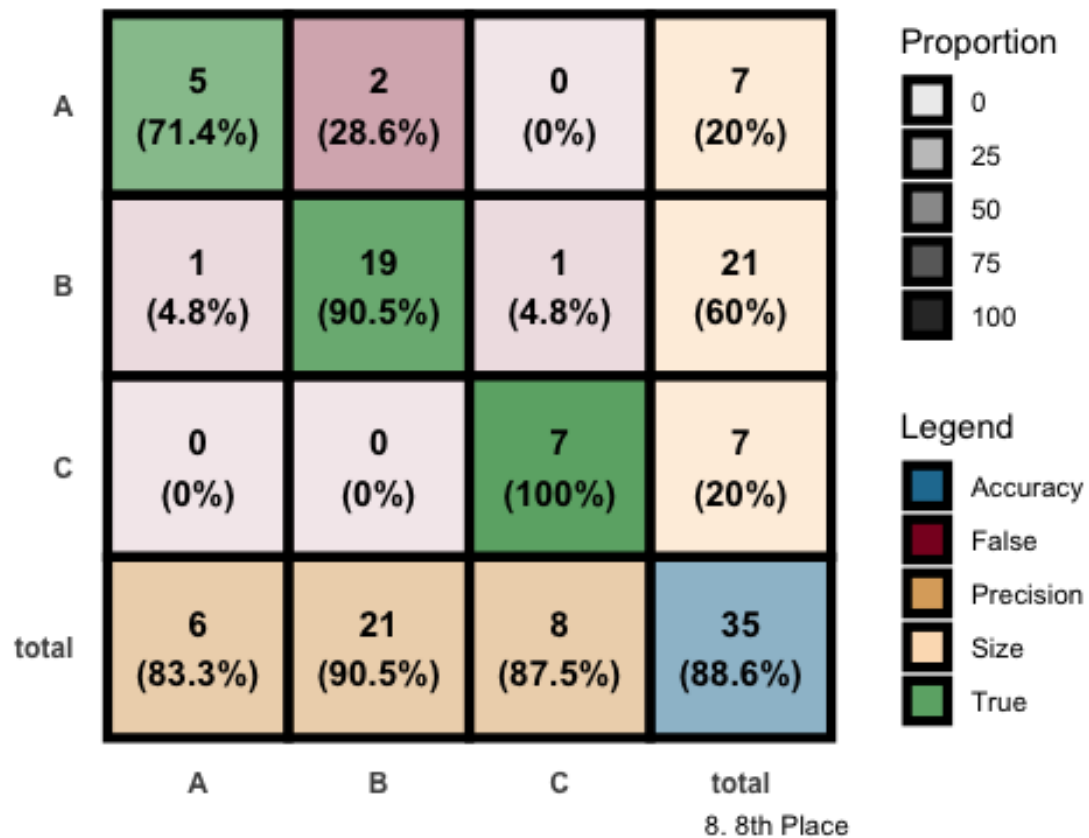
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
88.57%	0.627

Model Coefficients

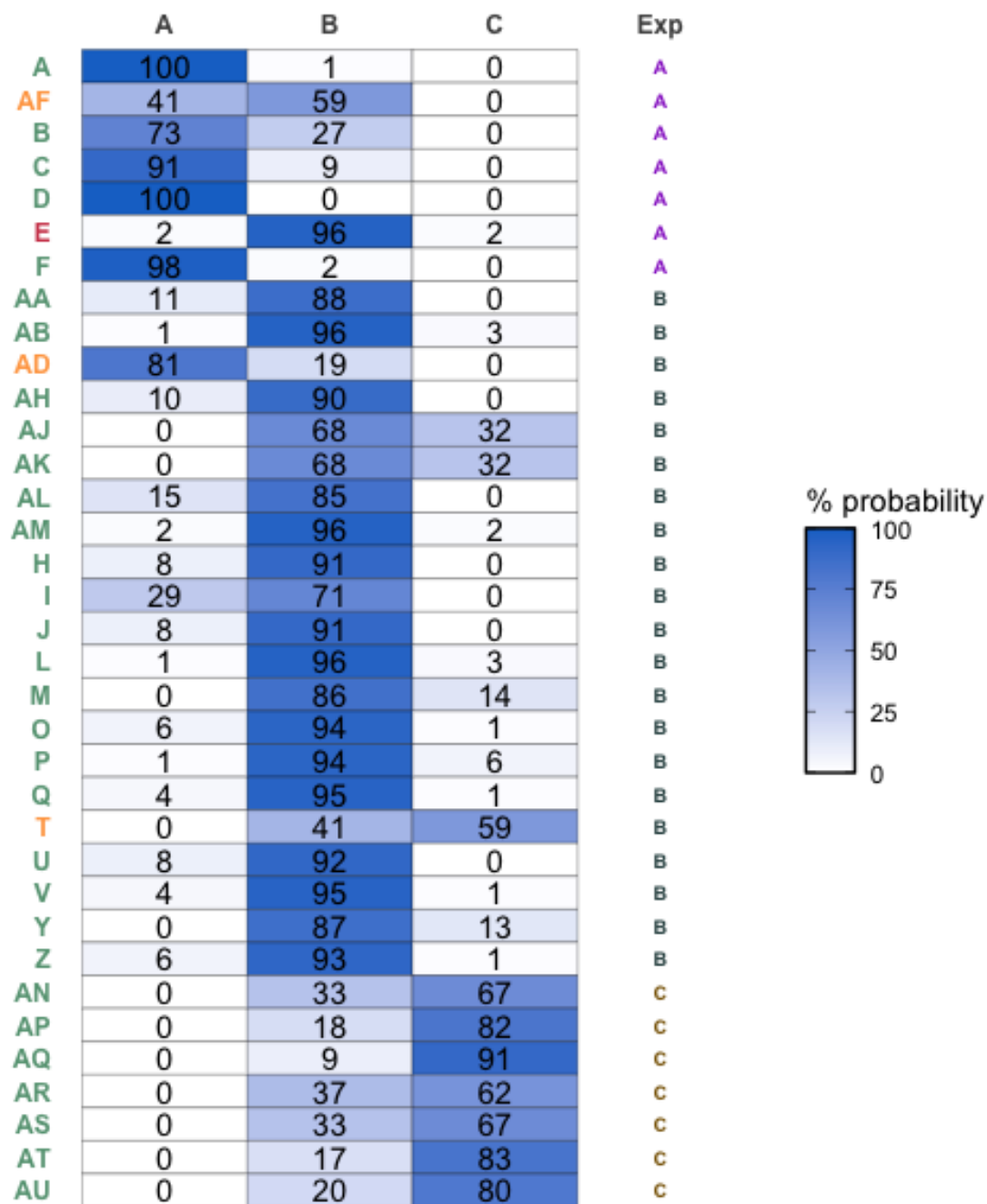
	x
NPA_1_P	91.0021318
Dist_(P,C)_P	-178.1867686
para_SM	0.1550555

Training Set Confusion Matrix



Training Set

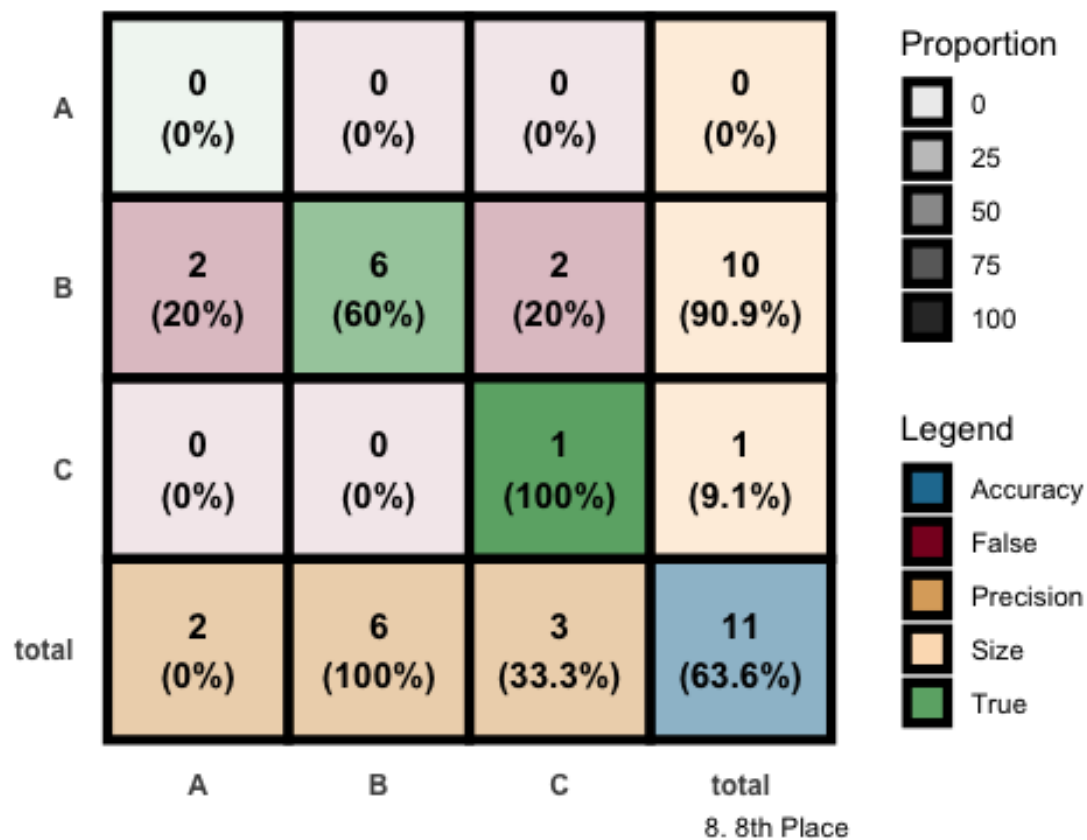
Probability Heatmap



8. 8th Place

Test Set

Confusion Matrix



Test Set

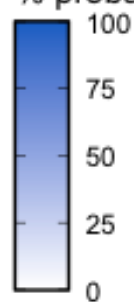
Probability Heatmap

	A	B	C
AC	1	95	4
AE	1	96	3
AI	0	20	80
G	0	45	55
K	3	96	1
N	78	22	0
R	3	96	1
S	100	0	0
W	0	70	29
X	2	96	2
AO	0	18	82

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



8. 8th Place

9th Place

Table:

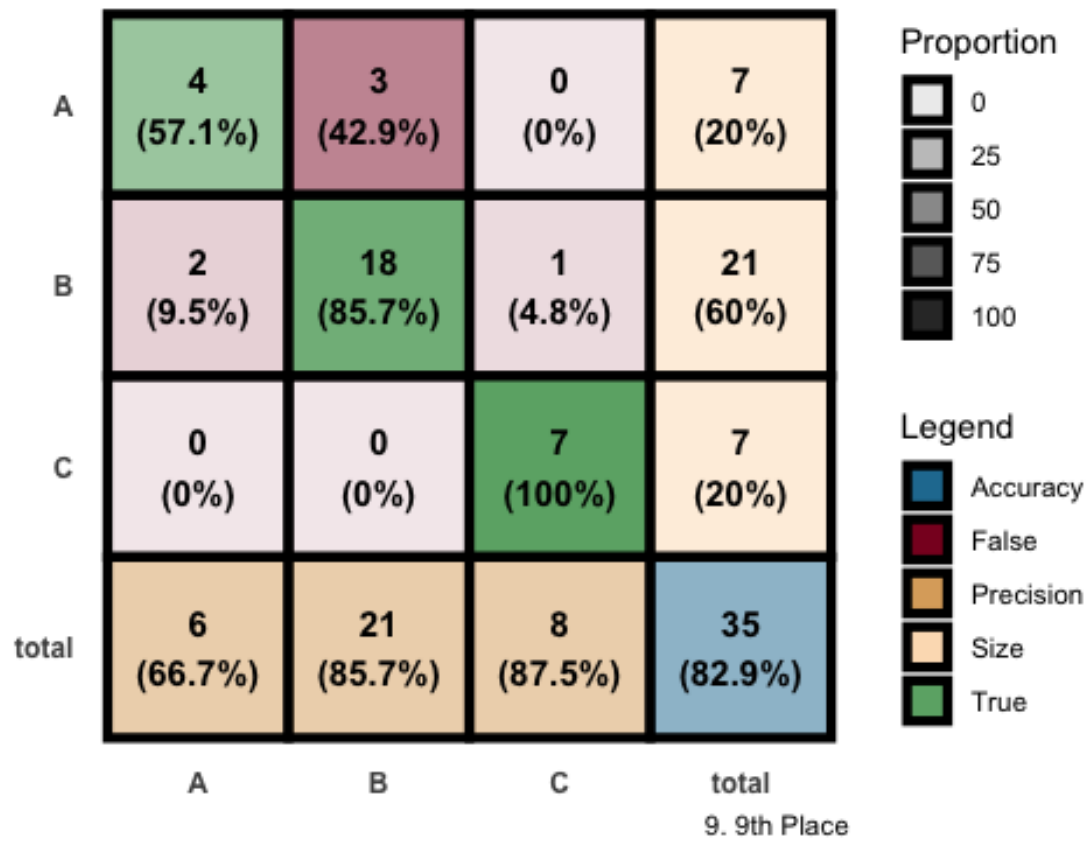
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
82.86%	0.626

Model Coefficients

	x
NPA_1_P	61.6879757
para_SM	0.1402902
fr_pyridine1	4.7039519
fr_pyridine2	0.2395840
fr_pyridine3	5.5632041

Training Set Confusion Matrix

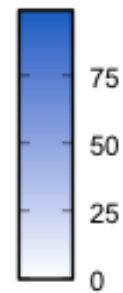


Training Set

Probability Heatmap

	A	B	C	Exp
A	74	26	0	A
AF	34	66	0	A
B	29	70	0	A
C	92	8	0	A
D	99	1	0	A
E	27	73	0	A
F	73	27	0	A
AA	1	89	10	B
AB	4	92	4	B
AD	24	75	0	B
AH	19	80	1	B
AJ	2	91	7	B
AK	1	81	18	B
AL	15	84	1	B
AM	3	92	5	B
H	9	90	2	B
I	26	73	0	B
J	2	91	7	B
L	6	92	2	B
M	4	92	4	B
O	0	57	42	B
P	65	35	0	B
Q	53	47	0	B
T	1	80	20	B
U	10	89	1	B
V	6	91	2	B
Y	0	41	59	B
Z	15	84	1	B
AN	0	39	61	C
AP	0	32	68	C
AQ	0	9	91	C
AR	0	31	69	C
AS	0	26	74	C
AT	0	19	81	C
AU	0	35	65	C

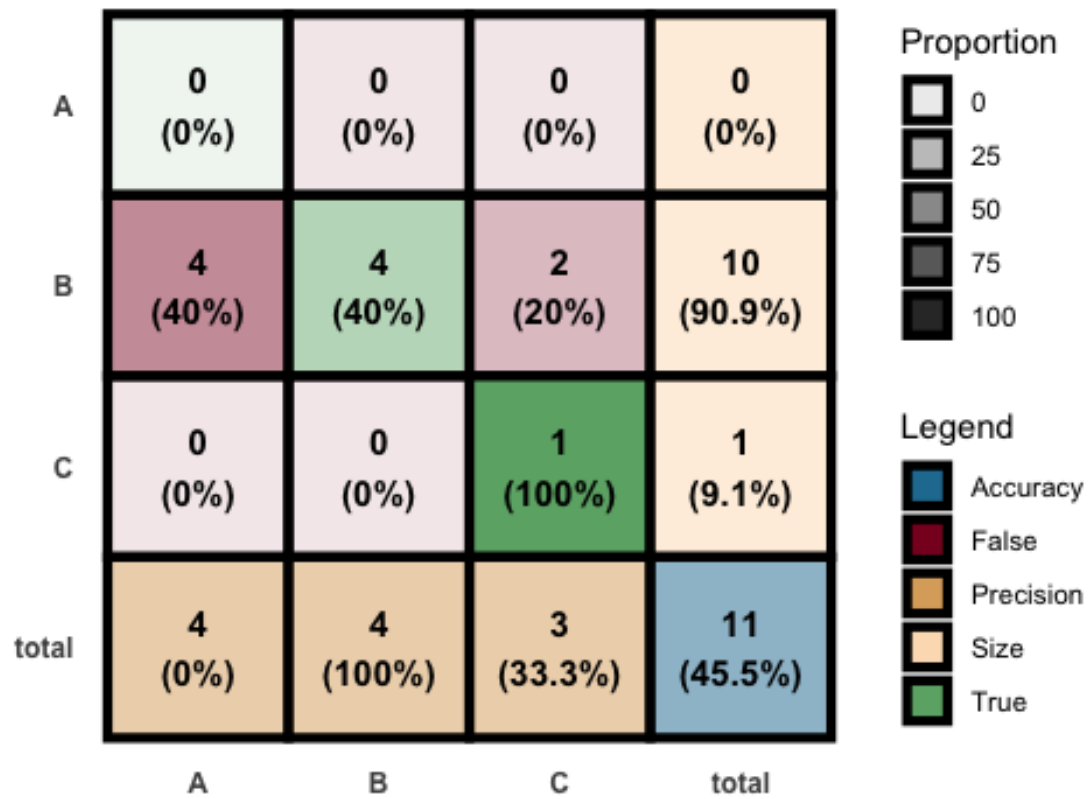
% probability



9. 9th Place

Test Set

Confusion Matrix



9. 9th Place

Test Set

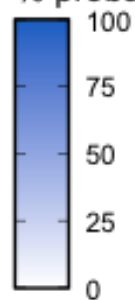
Probability Heatmap

	A	B	C
AC	8	90	2
AE	53	47	0
AI	20	80	1
G	0	18	82
K	1	89	9
N	99	1	0
R	3	92	5
S	100	0	0
W	0	49	51
X	73	27	0
AO	0	27	73

Exp

B
B
B
B
B
B
B
B
B
B
B
C

% probability



9. 9th Place

10th Place

Table:

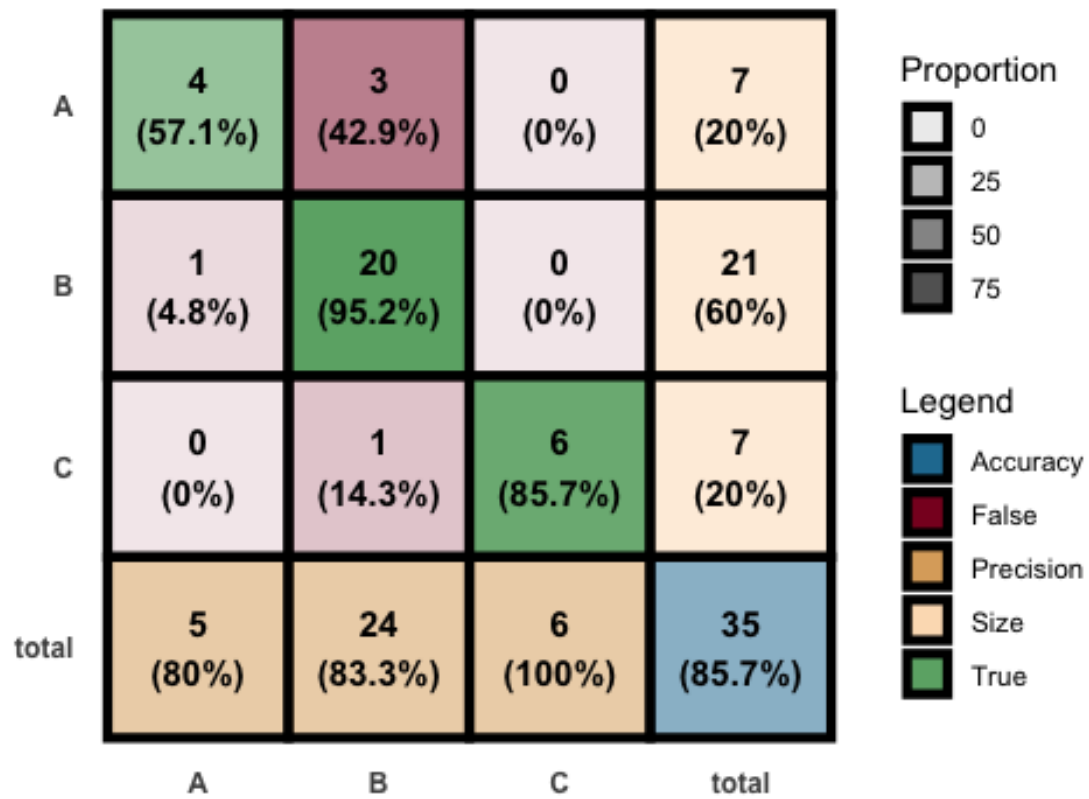
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
85.71%	0.623

Model Coefficients

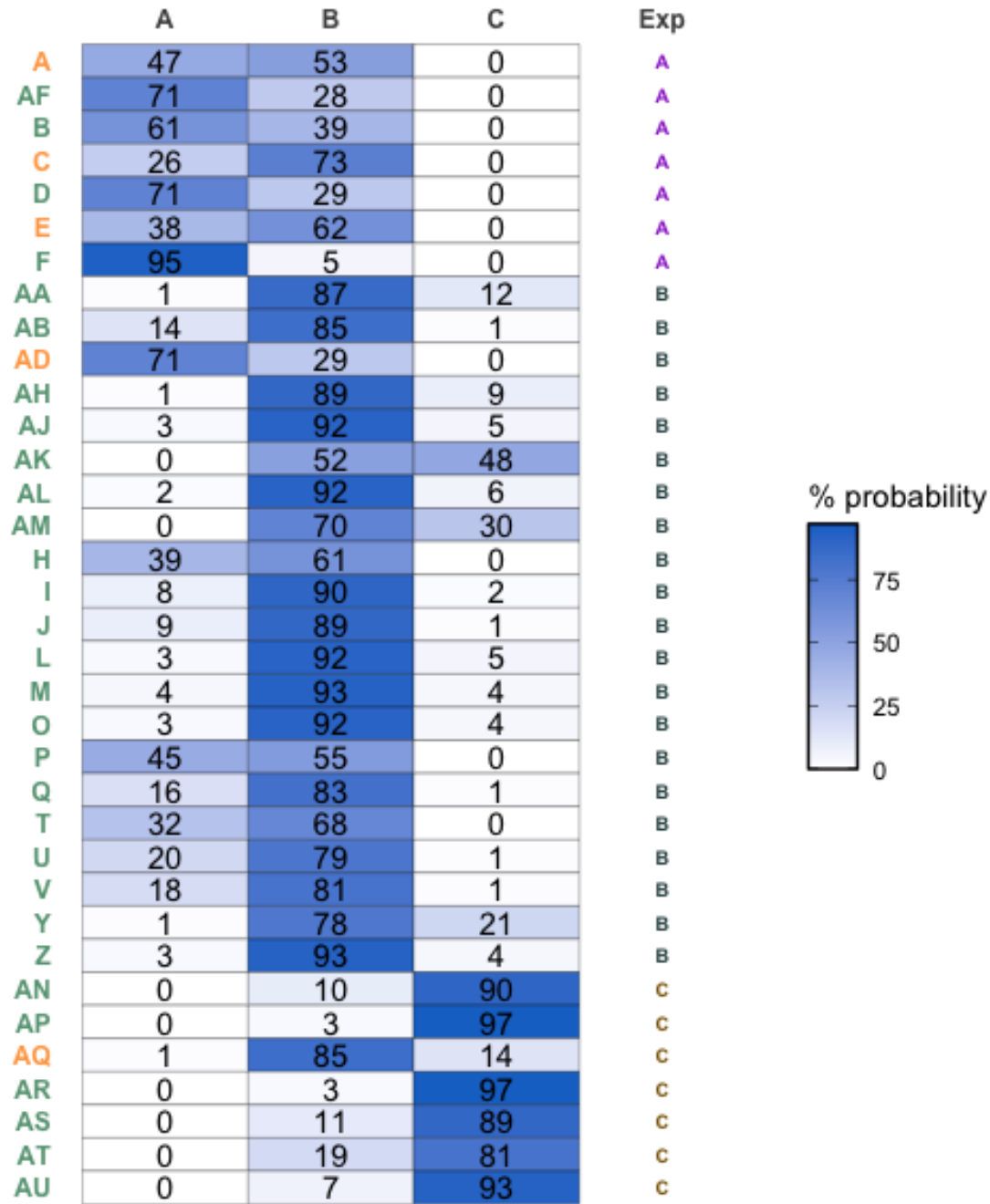
	x
NPA_6_SM	31.0312525
NPA_sum_SM	-11.1578676
para_SM	0.1060776

Training Set Confusion Matrix



Training Set

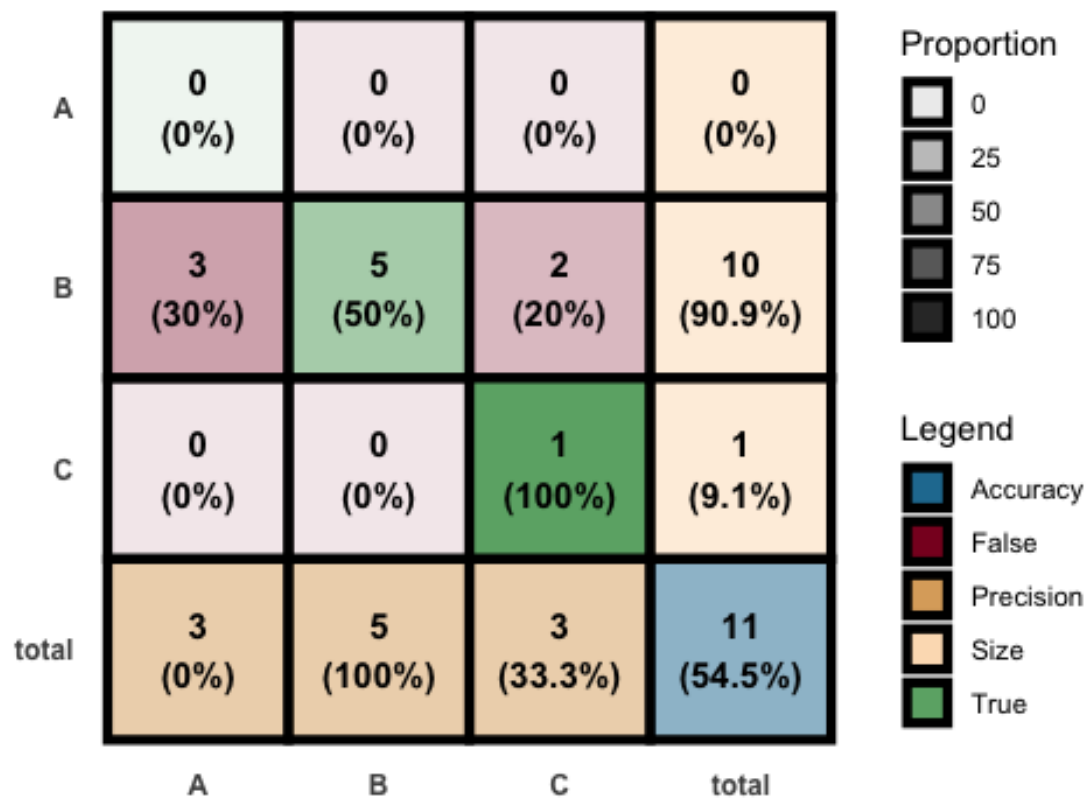
Probability Heatmap



10. 10th Place

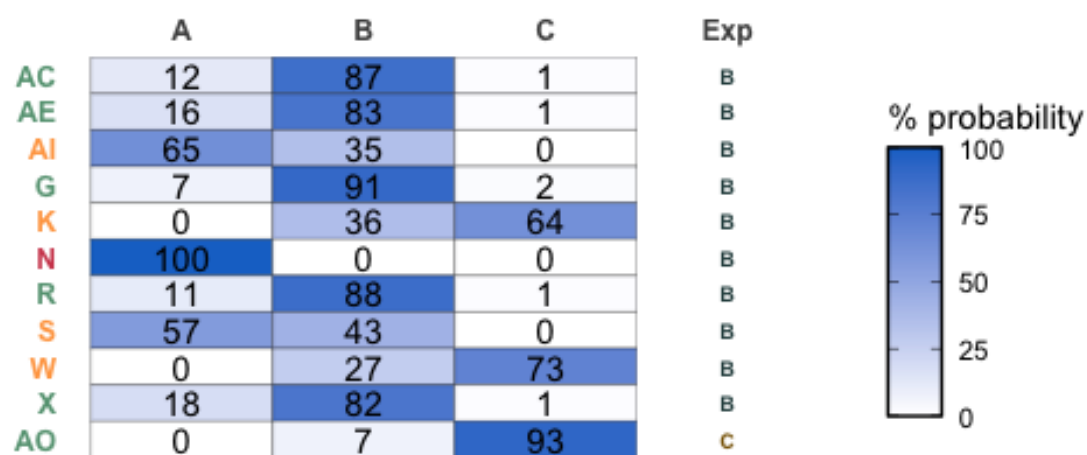
Test Set

Confusion Matrix



Test Set

Probability Heatmap



10. 10th Place