Pyridine Phosphination – 3 categories - Model Testing

Table of Contents

Model Search	
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	

Model Search

formula	McFadden R2
class \sim para_P + NPA_6_SM + Avg_NPA_SM	0.703
class \sim NPA_6_SM + para_SM + fr_C_O	0.662
class \sim NPA_6_SM + para_SM + fr_C_0_noCOO	0.662
class \sim Dist_(P,C)_P + para_SM + Iso.Polar_SM	0.651
class \sim para_P + NPA_6_SM + NPA_sum_SM	0.649
class \sim Dist_(P,C)_P + NPA_6_SM + para_SM	0.649
class \sim NPA_6_SM + Avg_NPA_SM + para_SM	0.646
$class \sim NPA_1_P + Dist_(P,C)_P + para_SM$	0.627
class \sim NPA_1_P + para_SM + fr_pyridine	0.626
class \sim 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

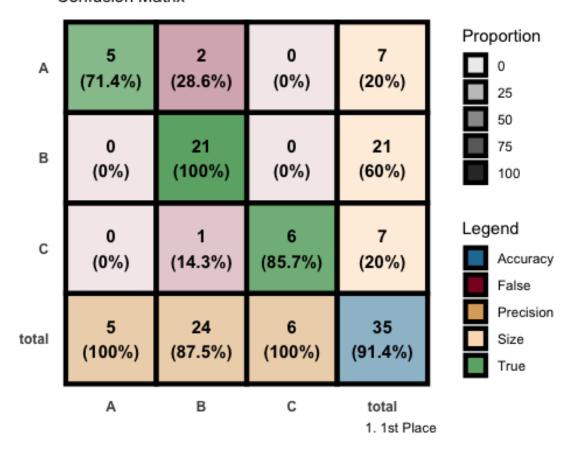
Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 91.43% 0.703

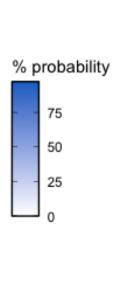
	X
para_P	0.1124591
NPA_6_SM	35.1047354
Avg_NPA_SM	-98.8554566

Training Set Confusion Matrix



Training Set Probability Heatmap

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

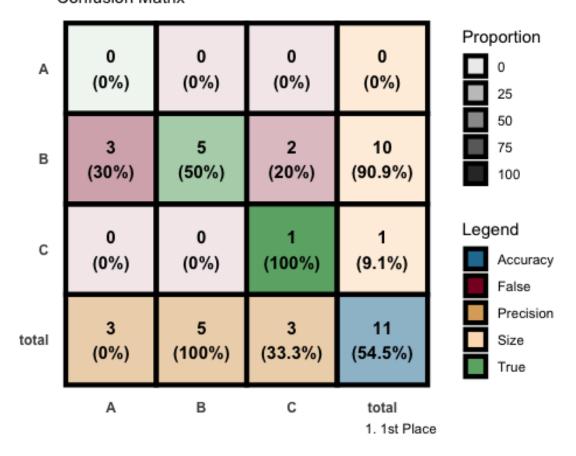


1. 1st Place

c c

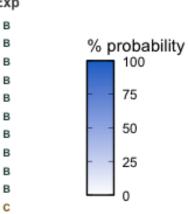
Ехр

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Exp
AC	14	85	1	В
AE	12	88	1	В
ΑI	73	27	0	В
G	2	94	4	В
K	0	42	58	В
N	100	0	0	В
R	40	60	0	В
S	86	14	0	В
W	0	26 75	74	В
Х	25	75	0	В
AO	0	11	89	С



1. 1st Place

2nd Place

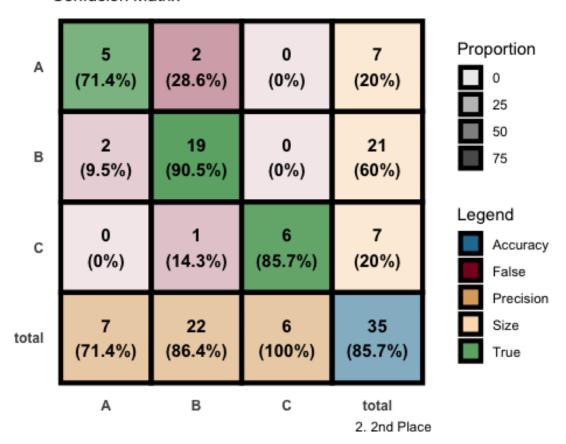
Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 85.71% 0.662

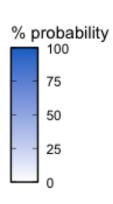
	X
NPA_6_SM	39.6636070
para_SM	0.1810645
fr_C_01	4.4955505
fr_C_02	17.9743375

Training Set Confusion Matrix



Training Set Probability Heatmap

	Α	В	С
Α	50	50	0
AF	97	3	0
В	84	16	0
С	16	84	0
D	80	20	0
E	46	54	0
F	82	18	0
AA	0	84	16
AB	7	95	1
AD	7	92	0
AH	7	92	0
AJ	10	89	0
AK	0	57	43
AL	3	96	1
AM	0	85	15
Н	64	36 77	0
- 1	23	77	0
J	1	96	3
L	4	95	1
M	6	94	0
0	0	94	6
Р	26	74	0
Q	1	97	7
Т	0	93	
U	2	97	0
V	83	17 91	
Υ	0	91	9
Z	0	67	33
AN	0	35	65
AP	0	12	88
AQ	0	64	36
AR	0	0	100
AS	0	0	100
AT	0	1	99 72
AU	0	28	72



2. 2nd Place

Exp

В

В

В

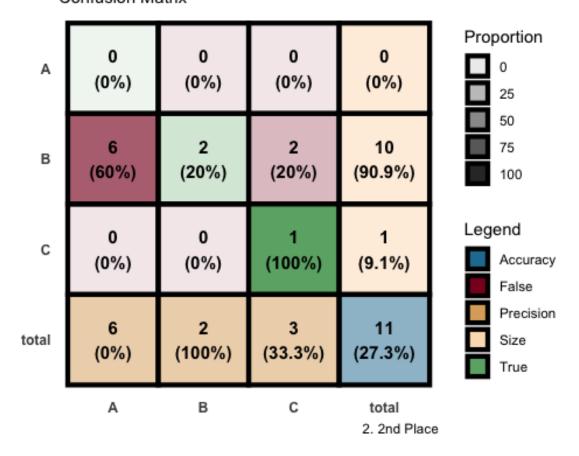
B B

В

B B

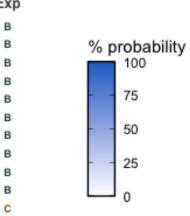
c c c

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Exp
AC	81	19	0	В
AE	0	35	65	В
ΑI	51	49	0	В
G	1	96	3	В
K	0	61	39	В
N	100	0	0	В
R	64	36	0	В
S	65	35	0	В
W	0	0	100	В
X	75	25	0	В
AO	0	34	66	С



2. 2nd Place

3rd Place

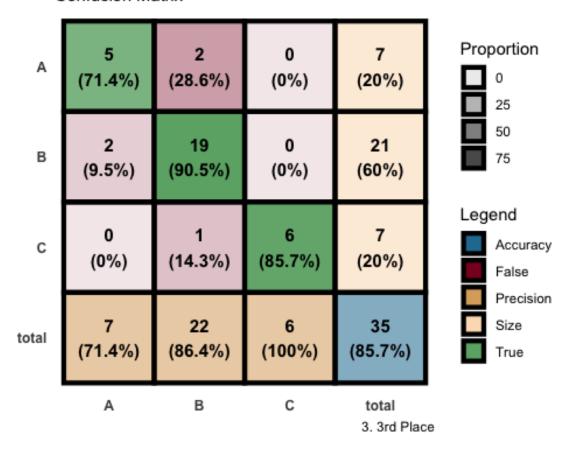
Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 85.71% 0.662

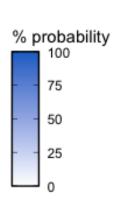
	X
NPA_6_SM	39.6636070
para_SM	0.1810645
fr_C_O_noCOO1	4.4955505
fr C O noCOO2	17.9743375

Training Set Confusion Matrix



Training Set Probability Heatmap

	Α	В	С
Α	50	50	0
AF	97	3	0
В	84	16	0
С	16	84	0
D	80	20	0
E	46	54	0
F	82	18	0
AA	0	84	16
AB	4	95	1
AD	7	92	0
AH	7	92	0
AJ	10	89	0
AK	0	57	43
AL	3	96	1
AM	0	85	15
Н	64	36 77	0
- 1	23	77	0
J	1	96	3
L	4	95	1
M	6	94	0
0	0	94	6
Р	26	74	0
Q	1	97	7
Т	0	93	
U	2	97	2
V	83	17	0
Υ	0	91	9
Z	0	67	33
AN	0	35	65
AP	0	12	88
AQ	0	64	36
AR	0	0	100
AS	0	0	100
AT	0	1	99 72
AU	0	28	72



3. 3rd Place

Exp

В

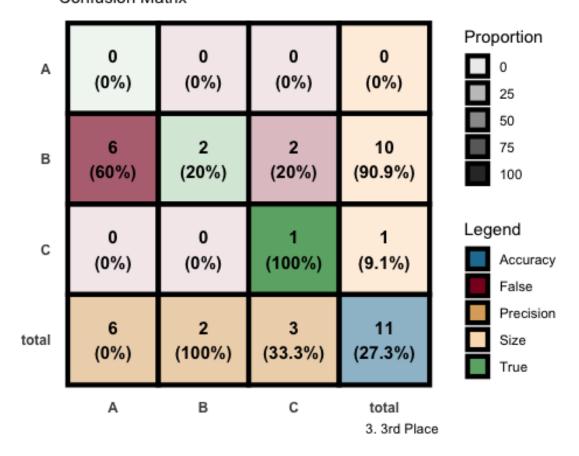
В

В

В

c c

Test Set Confusion Matrix



Test Set Probability Heatmap

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

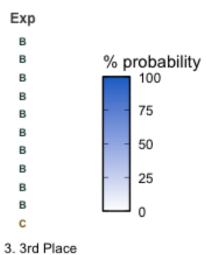


Table:

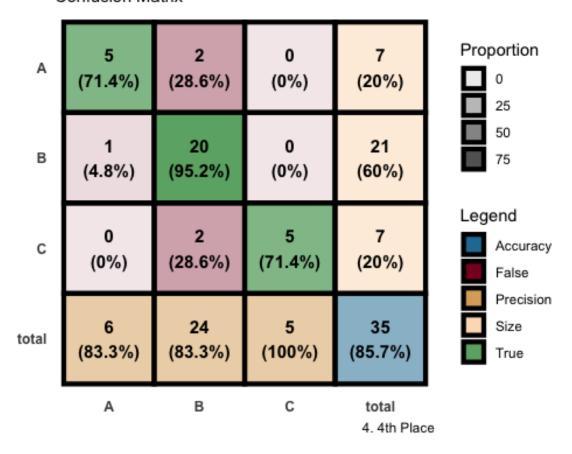
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 85.71% 0.651

Model Coefficients

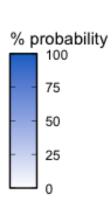
x
Dist_(P,C)_P -257.2488454
para_SM 0.1905674
Iso.Polar_SM 0.0329815

Training Set Confusion Matrix



Training Set Probability Heatmap

	Α	В	С
Α	40	60	0
AF	58	42	0
В	64	36	0
С	76	24	0
D	100	0	0
E	100 6	94	0
F	72	28	0
AA	18	82	0
AB	0 45	77	23
AD	45	55	0
AH	38	62	0
AJ	19	81	0
AK	0	96	4
AL	15	85	0
AM	4	96	0
Н	24	76	0
1.0	51	49	0
J	13	87	0
L	1	98	1
M	0	78	22
0	0	98	2
P	6	94	22 2 0
Q	4	96	0
Т	0	57	43
U	1	98	1
V	10	90	0
Υ	4	96	0
Z	0	88	12
AN	0	50	50
AP	0	3	97
AQ	0	20	80
AR	0	11 2 2 54	89
AS	0	2	98
AT	0	2	98
AU	0	54	46



4. 4th Place

Exp

В

В

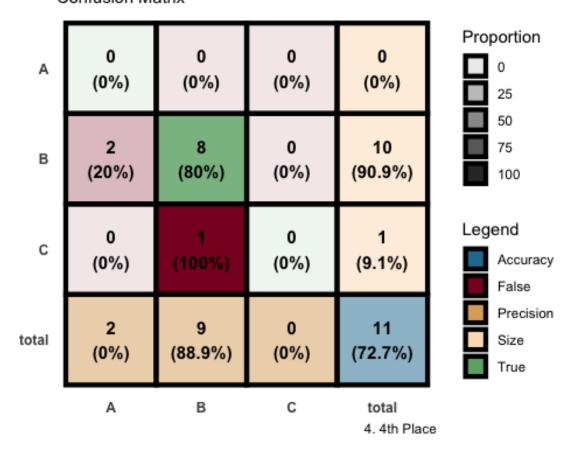
B B

В

B B

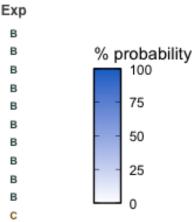
c c

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Ex
AC	1	98	1	В
AE	0	90	10	В
ΑI	0	73	27	В
G	8	92 88	0	В
K	12	88	0	В
N	98	2	0	В
R	9	91	0	В
S	100	0	0	В
W	1	98	1	В
X	0	97	3	В
AO	0	69	31	С



4. 4th Place

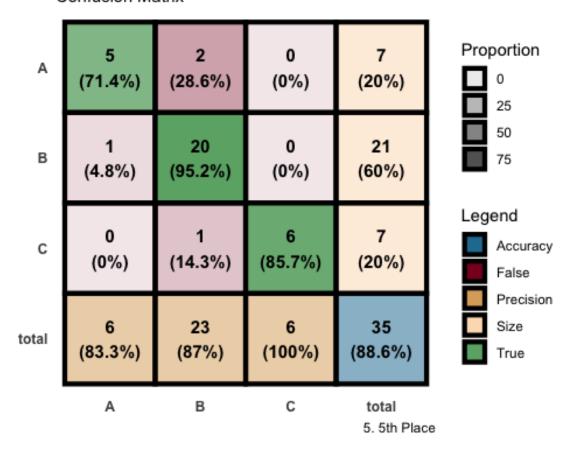
Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 88.57% 0.649

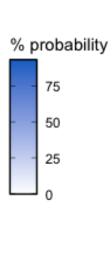
	X
para_P	0.0774822
NPA_6_SM	31.0525158
NPA_sum_SM	-13.5126258

Training Set Confusion Matrix



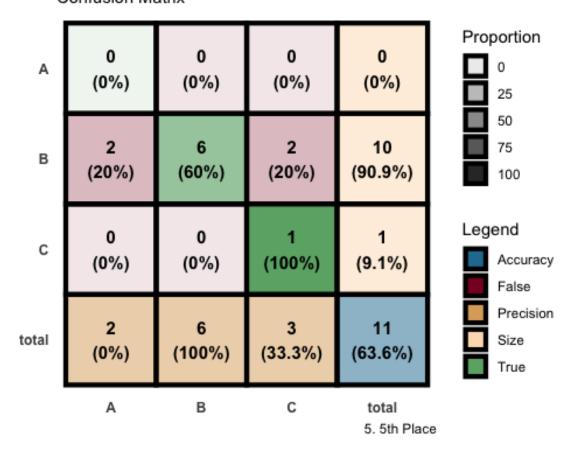
Training Set Probability Heatmap

	Α	В	С	Exp
Α	23	77	1	Α
AF	68	32	0	Α
В	51	49	0	Α
С	86	14	0	Α
D	83	17	0	Α
E	38	62	0	Α
F	85	15 77	0	Α
AA	1	77	23	В
AB	12	87		В
AD	23	87 76	1	В
AH	1 3	79	20	В
AJ	3	91	6	В
AK	0	67	32	В
AL	3	92	6	В
AM	0	73	27	В
Н	32	67	0	В
- 1	9	92 73 67 90	0 2 4	В
J	4	92	4	В
L	2	92 91	6	В
M	4	92	4	В
0	13	86	1	В
P	20	80	1	В
Q	35	65	0	В
T	61	39	0	В
U	16	83	1	В
V	13	86	1	В
Υ	1	87	12	В
Z	3	92	5	В
AN	0	40	60	С
AP	0	8	92	С
AQ	0	74	26	С
AR	0	8 7	92	С
AS	0	7	93	С
AT	0	13	87	С
ΑU	0	40	60	С



5. 5th Place

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Exp	
AC	10	88	1	В	
AE	14	85	1	В	% probability
ΑI	68	32	0	В	100
G	4	92	4	В	
K	0	39	61	В	75
N	100	0	0	В	50
R	27	73	0	В	50
S	43	56	0	В	25
W	0	23	77	В	25
Х	21	78	1	В	
AO	0	17	82	С	ŭ

5. 5th Place

Table:

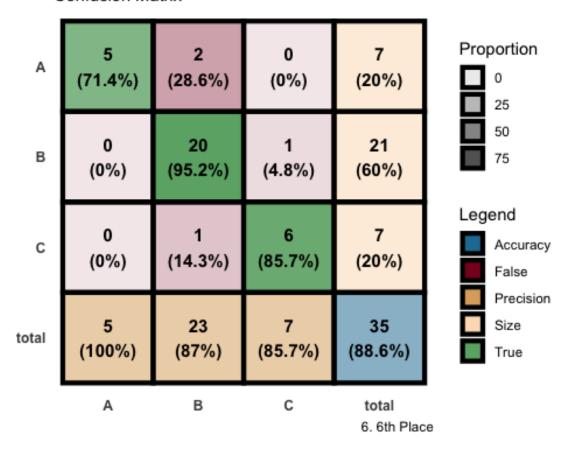
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 88.57% 0.649

Model Coefficients

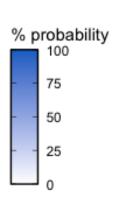
xDist_(P,C)_P-181.4237850NPA_6_SM33.1536447para_SM0.2071308

Training Set Confusion Matrix



Training Set Probability Heatmap

	Α	В	С
Α	40	60	0
AF	84	16	0
В	83	17	0
С	61	39	0
D	99	1	0
E	41	59	0
F	100	0	0
AA	18	0 82	0
AB	26 26	74	0
AD	26	74	0
AH	2	98	1
AJ	21	79	0
AK	0	14	86
AL	1	98	1
AM	0	84	16
Н	20	80	0
- 1	15	85	0
J	1	98	1
L	0	95	5
M	0	90	10
0	0	97	2
P	3	96	0
Q	1	98	2
Т	0	81	19
U	13	86	0
V	38	62	0
Υ	2	97	1
Z	1	98	1
AN	0	8	92
AP	0	2	98
AQ	0	97	3
AR	0	6	94
AS	0	20	80
AT	0	10	90
AU	0	4	96



6. 6th Place

Ехр

B B

В

В

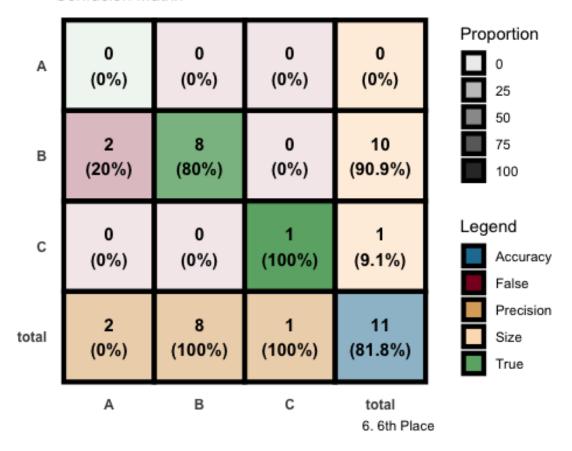
B B

В

B B

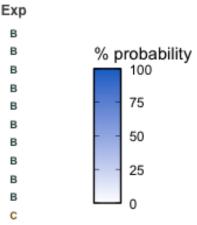
c c

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	E
AC	12	88	0	
AE	0	96	4	
ΑI	1	98	2	
G	38	62	0	
K	0	83	17	
N	100 23	0	0	
R S	23	77	0	
S	80	20 88 76	0	
W	0	88	12	
Х	0 24	76	0	
AO	0	6	94	



6. 6th Place

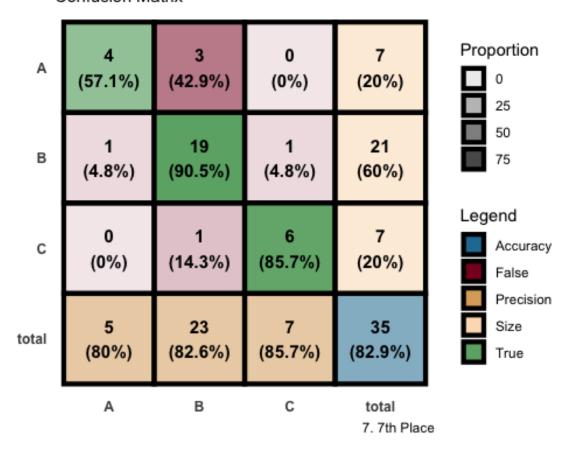
Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 82.86% 0.646

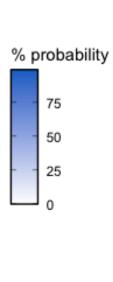
	X
NPA_6_SM	37.2573966
Avg_NPA_SM	-69.5656894
para_SM	0.1550816

Training Set Confusion Matrix



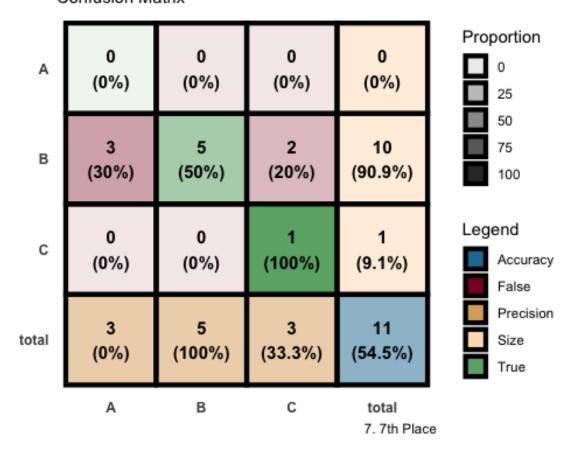
Training Set Probability Heatmap

	Α	В	С	Exp
Α	45	55	0	A
AF	85	15	0	A
В	73	27	0	A
C	21	78	0	A
D	81	19	0	A
E	34	65	0	A
F	99	1	0	A
AA	1	92	8	В
AB	18	82	0	В
AD	72	28	0	В
AH	1	95	4	В
AJ	1	95	4	В
AK	0	41	59	В
AL	2	95	3	В
AM	0	71	29	В
Н	44	55	0	В
- 1	10	89	1	В
J	5	94	1	В
L	2 2	95	3 2 7	В
M	2	95	2	В
0	1	92		В
P	45	55	0	В
Q	8 7	92	1	В
Т	7	92	1	В
U	23	77	0	В
V	24	76	0	В
Υ	0	81	19	В
Z	3	95	2	В
AN	0	4	96	С
AP	0	1	99	С
AQ	0	86	14	С
AR	0	1	99	С
AS	0	5	95	С
AT	0	10	90	С
AU	0	3	97	С



7. 7th Place

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Exp	
AC	17	82	0	В	
AE	8	91	1	В	% probability
ΑI	63	37	0	В	100
G	6	93	1	В	
K	0	29	71	В	75
N	100	0	0	В	
R	11	88	0	В	50
S	83	17	0	В	25
W	0	24	76	В	[] 25
X	20	80	0	В	
AO	0	3	97	С	ŭ

7. 7th Place

8th Place

Table:

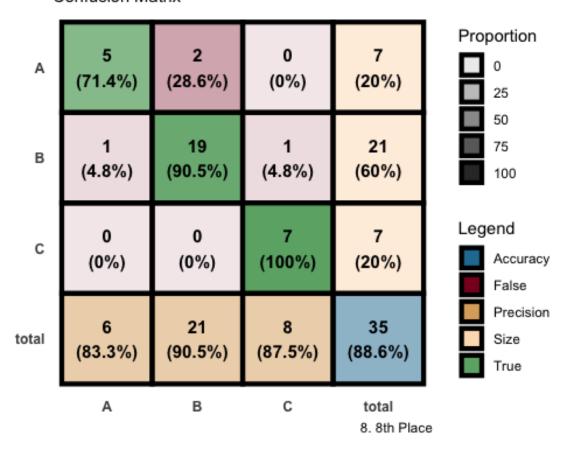
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 88.57% 0.627

Model Coefficients

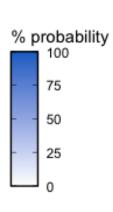
	X
NPA_1_P	91.0021318
Dist_(P,C)_P	-178.1867686
para_SM	0.1550555

Training Set Confusion Matrix



Training Set Probability Heatmap

	Α	В	С
Α	100	1	0
AF	41	59	0
В	73	27	0
C	91	9	0
D	100	0	0
E	100 2	96	2
F	98	2 88	0
AA	11	88	0
AB	1 81	96	3
AD	81	19	0
AH	10	90	0
AJ	0	68	32
AK	0	68	32
AL	15	85	0
AM	2	96	2
Н	8	91 71	0
- 1	29	71	0
J	8	91	0
L	1	96	3
M	0	86	14
0	6	94	1
P	1	94	6
Q	4	95	1
T	0	41	59
U	8	92	0
V	4	95	1
Υ	0	87	13
Z	6	93	1
AN	0	33	67
AP	0	18	82
AQ	0	9	91
AR	0	37	62
AS	0	33	67
AT	0	17 20	83
AU	0	20	80



8. 8th Place

Exp

В

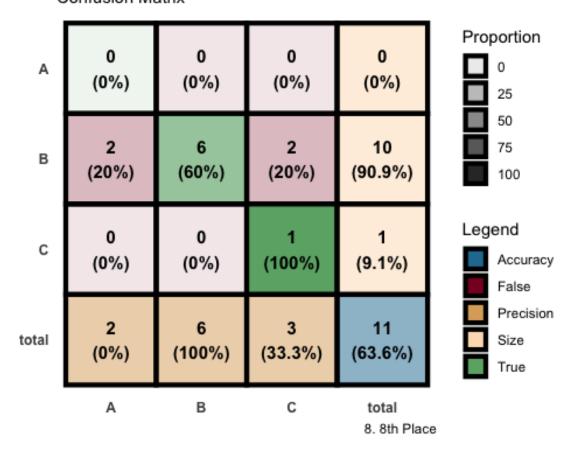
В

В

B B

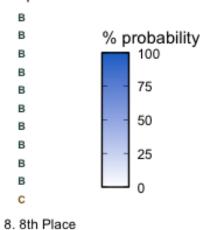
c c

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Exp
AC	1	95	4	В
AE	1	96	3	В
ΑI	0	20	80	В
G	0	45	55	В
K	3	96	1	В
N	78	22 96	0	В
R	3	96	1	В
S	100	0	0	В
W	0	70	29	В
Х	2	96	2	В
AO	0	18	82	С



9th Place

Table:

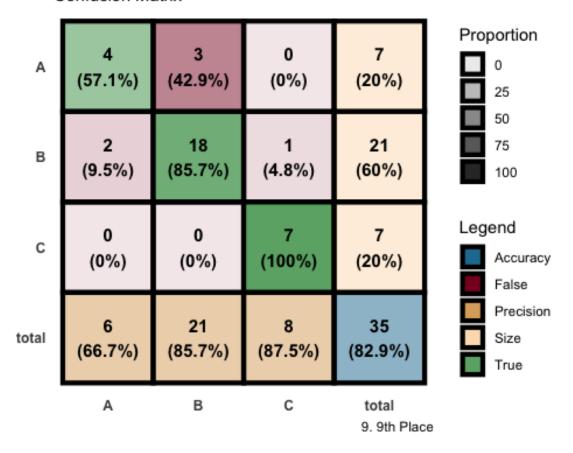
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2		
82.86%	0.626		

Model Coefficients

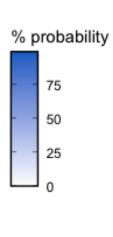
	X
NPA_1_P	61.6879757
para_SM	0.1402902
fr_pyridine1	4.7039519
fr_pyridine2	0.2395840
fr_pyridine3	5.5632041

Training Set Confusion Matrix



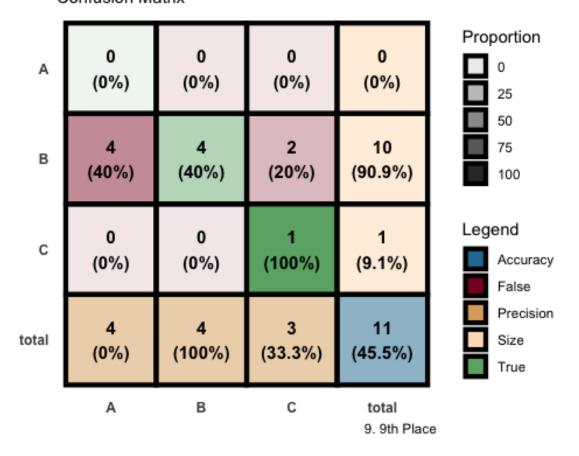
Training Set Probability Heatmap

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



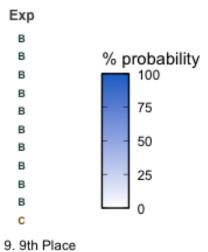
9. 9th Place

Test Set Confusion Matrix



Test Set Probability Heatmap

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



10th Place

Table:

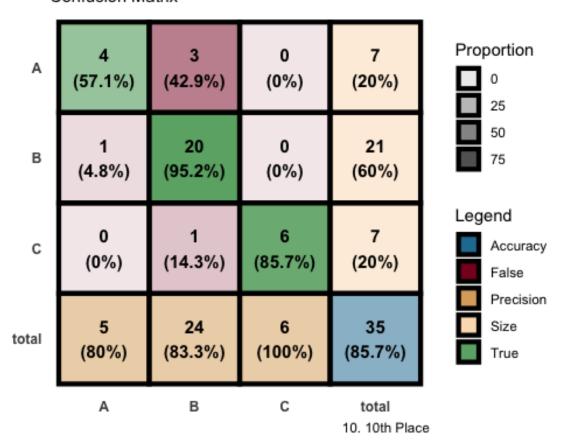
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy McFadden_R2 85.71% 0.623

Model Coefficients

	X
NPA_6_SM	31.0312525
NPA_sum_SM	-11.1578676
para_SM	0.1060776

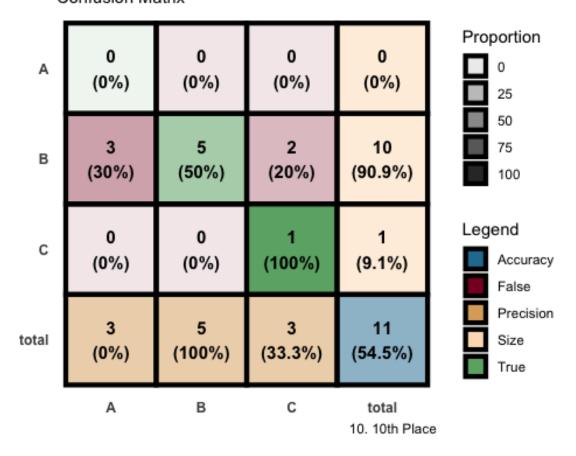
Training Set Confusion Matrix



Training Set Probability Heatmap

	Α	В	С	Exp	
Α	47	53	0	A	
AF	71	28	0	Α	
В	61	39	0	A	
С	26	73	0	Α	
D	71	29	0	A	
E	38	62	0	A	
F	95	5	0	Α	
AA	1	87	12	В	
AB	14	85	1	В	
AD	71	29	0	В	
AH	1	89	9	В	
AJ	3	92	5	В	
AK	0	52	48	В	
AL	2	92	6	В	% probability
AM	0	70	30	В	
Н	39	61	0	В	
- 1	8	90	2	В	75
J	9	89	1	В	50
L	3	92	5	В] 50
M	4	93	4	В	- 25
0	3	92	4	В	23
Р	45	55	0	В	
Q	16	83	1	В	v
Т	32	68	0	В	
U	20	79	1	В	
V	18	81	1	В	
Υ	1	78	21	В	
Z	3	93	4	В	
AN	0	10	90	С	
AP	0	3	97	С	
AQ	1	85	14	С	
AR	0	3	97	С	
AS	0	11	89	С	
AT	0	19	81	С	
AU	0	7	93	С	

Test Set Confusion Matrix



Test Set Probability Heatmap

	Α	В	С	Exp	
AC	12	87	1	В	
AE	16	83	1	В	% probability
ΑI	65	35	0	В	100
G	7	91	2	В	
K	0	36	64	В	75
N	100	0	0	В	50
R	11	88	1	В	50
S	57	43	0	В	25
W	0	27	73	В	25
Х	18	82	1	В	
AO	0	7	93	С	ŭ

10. 10th Place