

## Pyr Phos - Binary - works against not

### Table of Contents

Section Results Summary .....	1
Model Search.....	1
1st Place.....	2
2nd Place.....	7
3rd Place.....	12
4th Place.....	17
5th Place.....	22
5th Place with 5 features (fr_halogen added) .....	27
6th Place.....	32
7th Place.....	37
8th Place.....	42
9th Place.....	47
10th Place.....	52

### Section Results Summary

	Training Accuracy	Test Accuracy
1. 1st Place	82.61	68.8
2. 2nd Place	84.78	87.5
3. 3rd Place	84.78	68.8
4. 4th Place	89.13	81.2
5. 5th Place	89.13	81.2
5. 5th Place with 5 features	91.30	87.5
6. 6th Place	82.61	93.8
7. 7th Place	86.96	81.2
8. 8th Place	86.96	81.2
9. 9th Place	84.78	75.0
10. 10th Place	86.96	87.5

### Model Search

formula	McFadden R2
class ~ cross.angle_P + para.angle_P + fr_Ar_N + fr_bicyclic	0.526
class ~ fr_Ar_N + fr_aryl_methyl + fr_benzene + fr_halogen	0.520
class ~ para_P + Dist(Hhal,C)_SM + fr_Ar_N + fr_amide	0.518
class ~ NPA_sum_SM + fr_Ar_N + fr_aryl_methyl + fr_benzene	0.512
class ~ Avg_NPA_SM + fr_Ar_N + fr_aryl_methyl + fr_benzene	0.512
class ~ NPA_5_P + fr_Ar_N + fr_aryl_methyl + fr_benzene	0.509

formula	McFadden R2
class ~ NPA_6_SM + fr_Ar_N + fr_aryl_methyl + fr_benzene	0.504
class ~ Avg_NPA_SM + Dist(Hhal,C)_SM + fr_Ar_N + fr_aryl_methyl	0.490
class ~ para.angle_P + NPA_6_SM + fr_Ar_N + fr_aryl_methyl	0.483
class ~ NPA_sum_SM + Dist(Hhal,C)_SM + fr_Ar_N + fr_aryl_methyl	0.483

### 1st Place

Table:

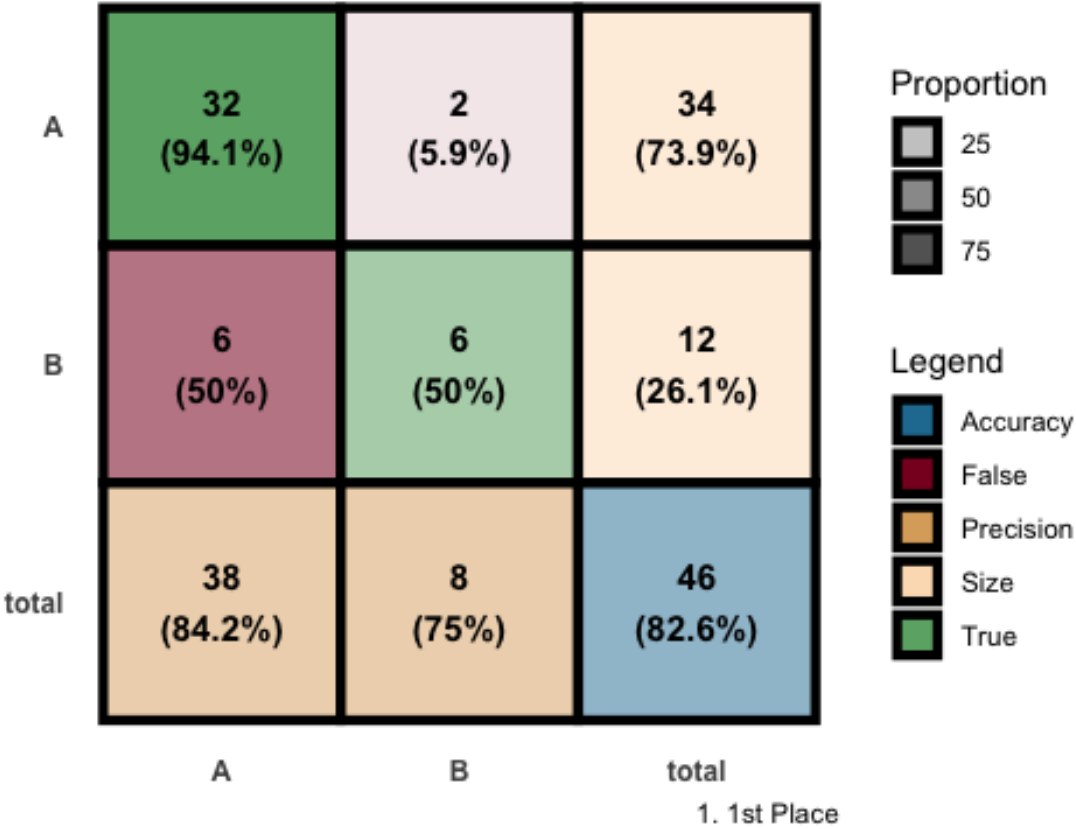
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
82.61%	0.526

Model Coefficients

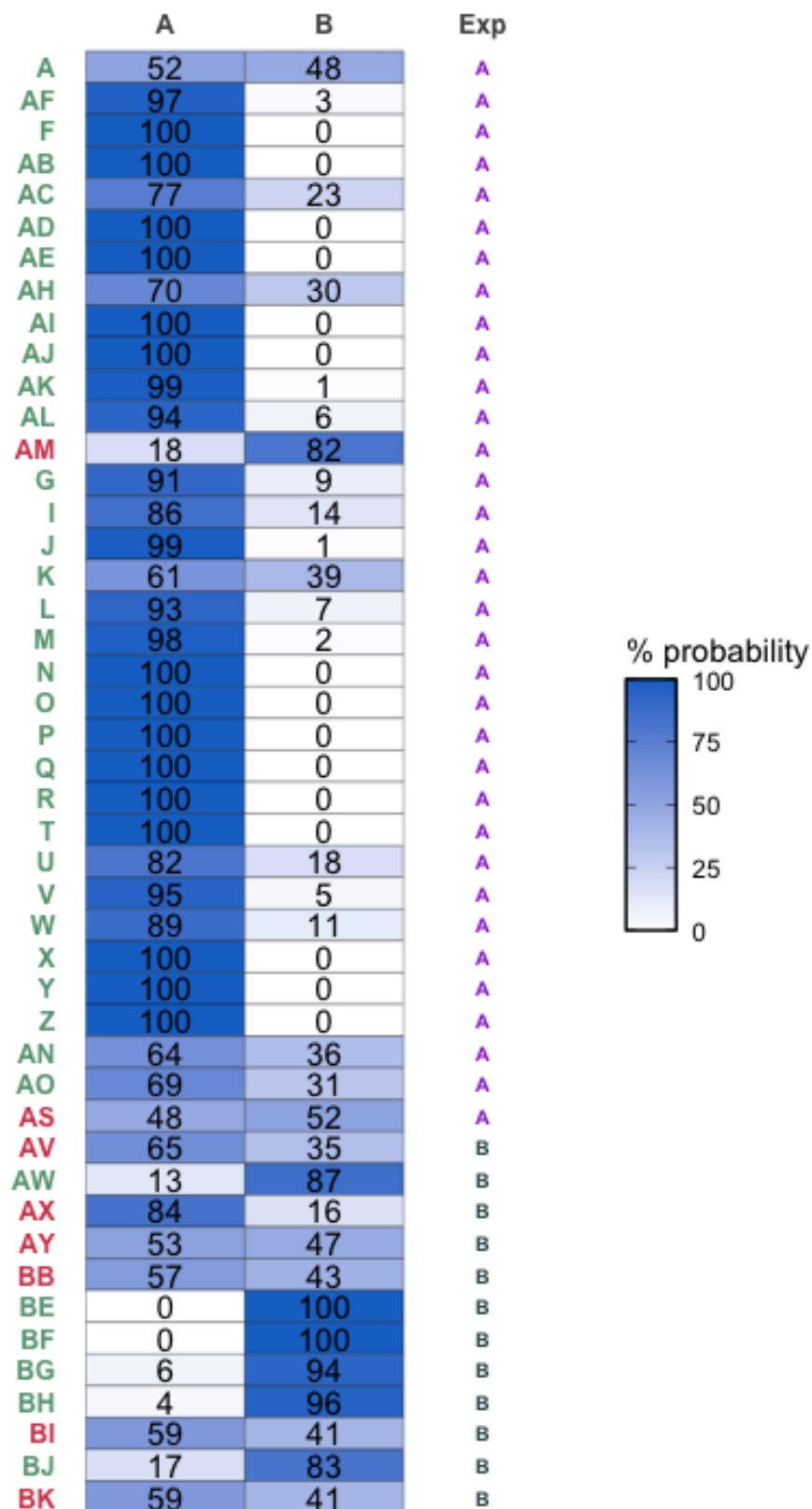
	x
(Intercept)	90.574398
cross.angle_P	-0.490975
para.angle_P	-0.530693
fr_Ar_N1	-46.303760
fr_Ar_N2	-96.040002
fr_Ar_N3	-89.766140
fr_bicyclic1	-4.672925
fr_bicyclic2	-69.526693
fr_bicyclic3	-55.854458
fr_bicyclic5	-53.824196
fr_bicyclic6	-11.228785

Training Set  
Confusion Matrix



# Training Set

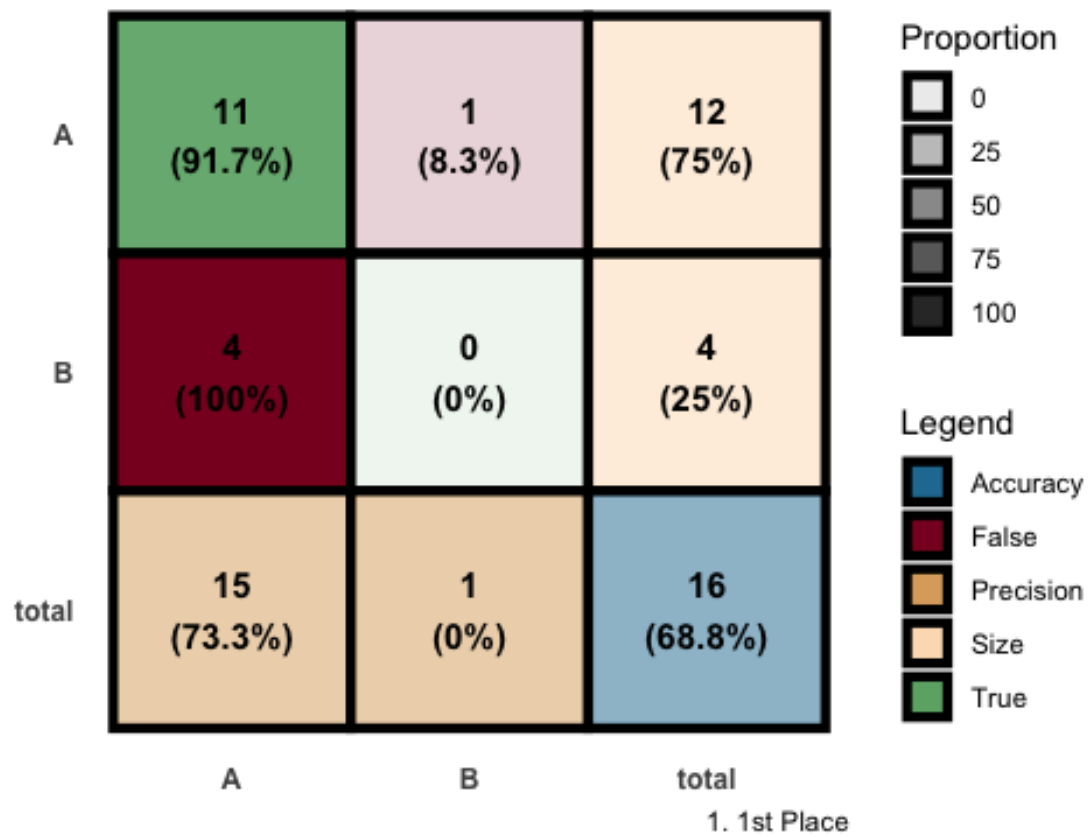
## Probability Heatmap



1. 1st Place

## Test Set

### Confusion Matrix



## Test Set

### Probability Heatmap

	A	B
B	91	9
C	100	0
D	100	0
E	100	0
AA	76	24
H	9	91
S	100	0
AP	51	49
AQ	84	16
AR	70	30
AT	100	0
AU	61	39
AZ	55	45
BA	84	16
BC	84	16
BD	99	1

Exp

A

A

A

A

A

A

A

A

A

A

A

A

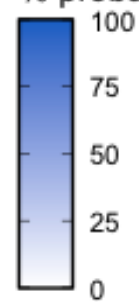
B

B

B

B

% probability



1. 1st Place

## 2nd Place

Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
84.78%	0.52

Model Coefficients

	x
(Intercept)	14.653960
fr_Ar_N1	-14.641780
fr_Ar_N2	-28.886527
fr_Ar_N3	-30.760596
fr_aryl_methyl1	-3.720066
fr_aryl_methyl2	-16.691481
fr_benzene1	-2.364301
fr_benzene2	-35.077472
fr_halogen1	2.342569
fr_halogen2	4.212291
fr_halogen3	1.848104
fr_halogen4	-1.081430
fr_halogen5	2.352174
fr_halogen6	-9.373352

# Training Set Confusion Matrix

A	<div>30</div> <div>(88.2%)</div>	<div>4</div> <div>(11.8%)</div>	<div>34</div> <div>(73.9%)</div>
B	<div>3</div> <div>(25%)</div>	<div>9</div> <div>(75%)</div>	<div>12</div> <div>(26.1%)</div>
total	<div>33</div> <div>(90.9%)</div>	<div>13</div> <div>(69.2%)</div>	<div>46</div> <div>(84.8%)</div>
	A	B	total

## Legend

- Accuracy
- False
- Precision
- Size
- True

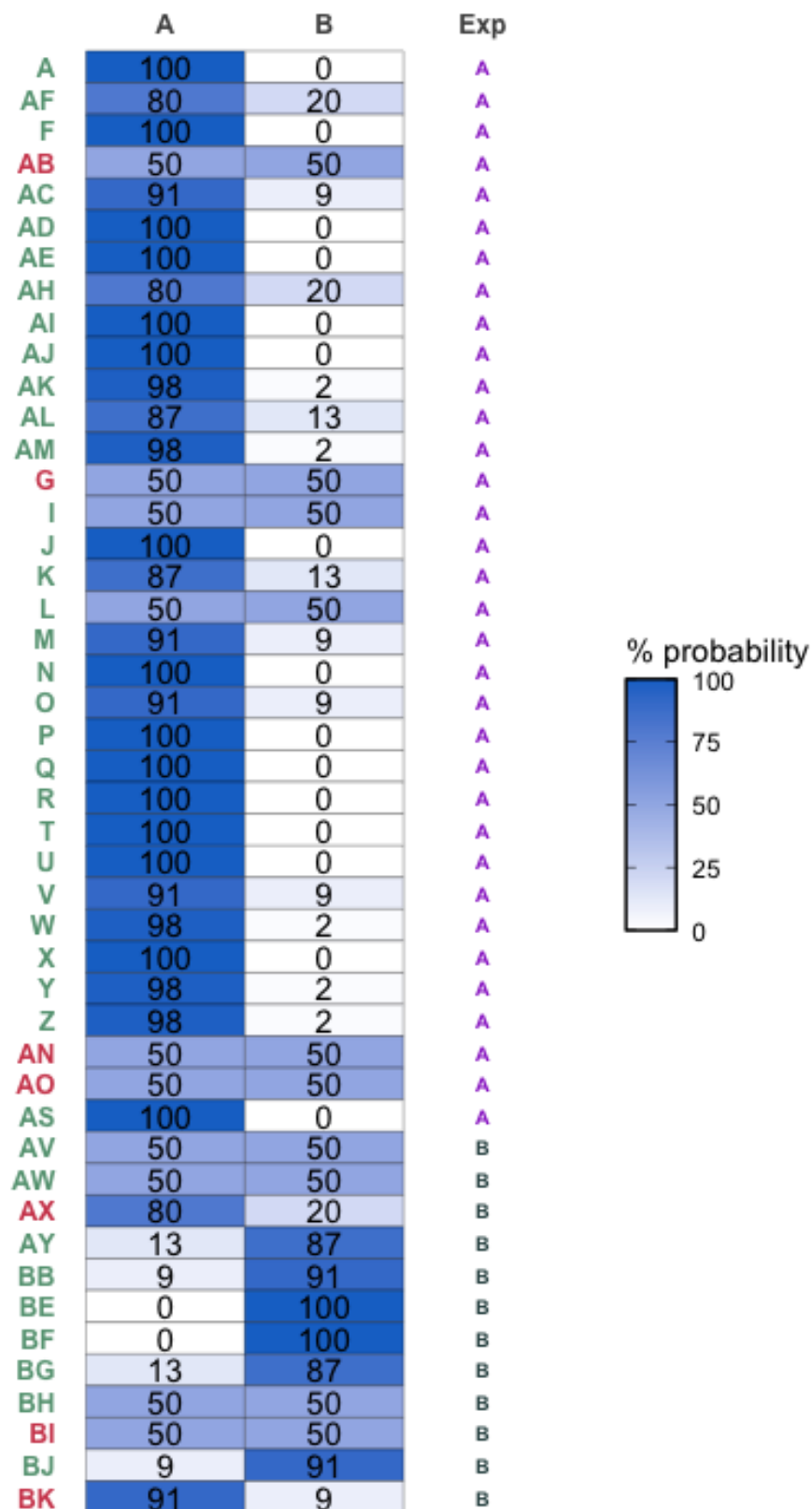
## Proportion

- 20
- 40
- 60
- 80



## Training Set

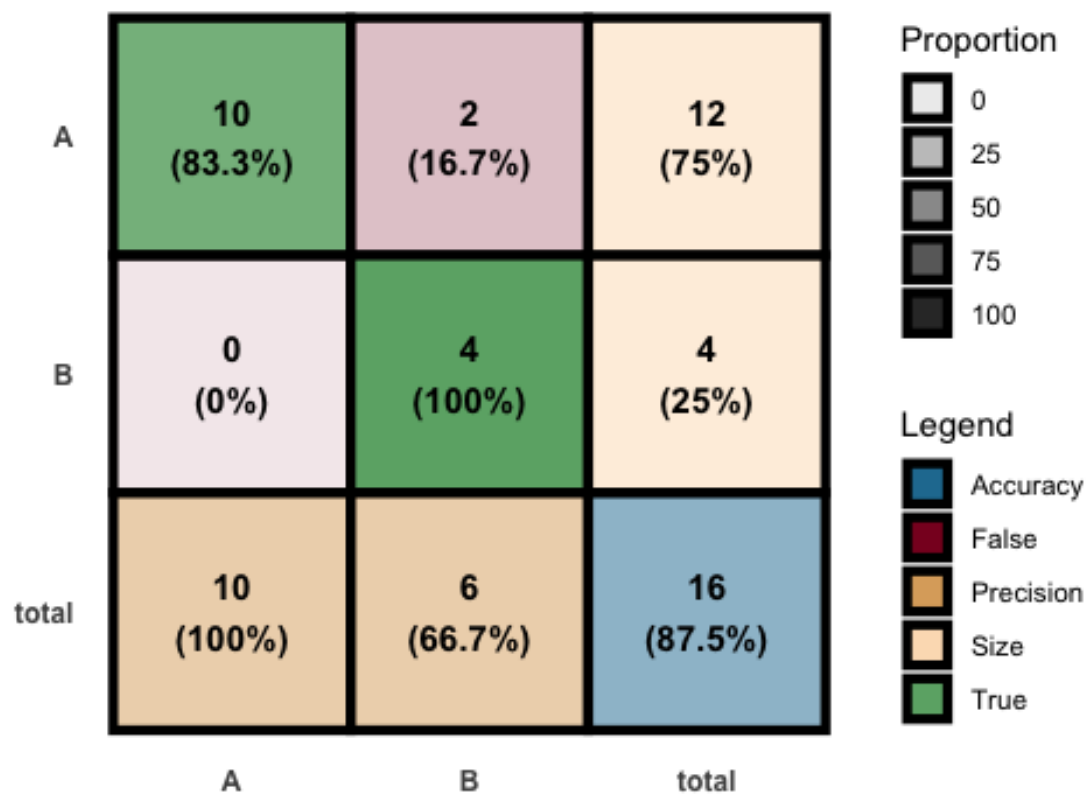
### Probability Heatmap



2. 2nd Place

## Test Set

### Confusion Matrix



2. 2nd Place

## Test Set

### Probability Heatmap

	A	B
B	91	9
C	100	0
D	100	0
E	100	0
AA	50	50
H	50	50
S	100	0
AP	91	9
AQ	50	50
AR	100	0
AT	100	0
AU	91	9
AZ	50	50
BA	13	87
BC	50	50
BD	50	50

Exp

A

A

A

A

A

A

A

A

A

A

A

A

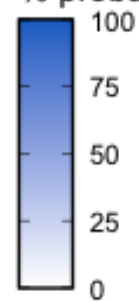
B

B

B

B

% probability



2. 2nd Place

### 3rd Place

Table:

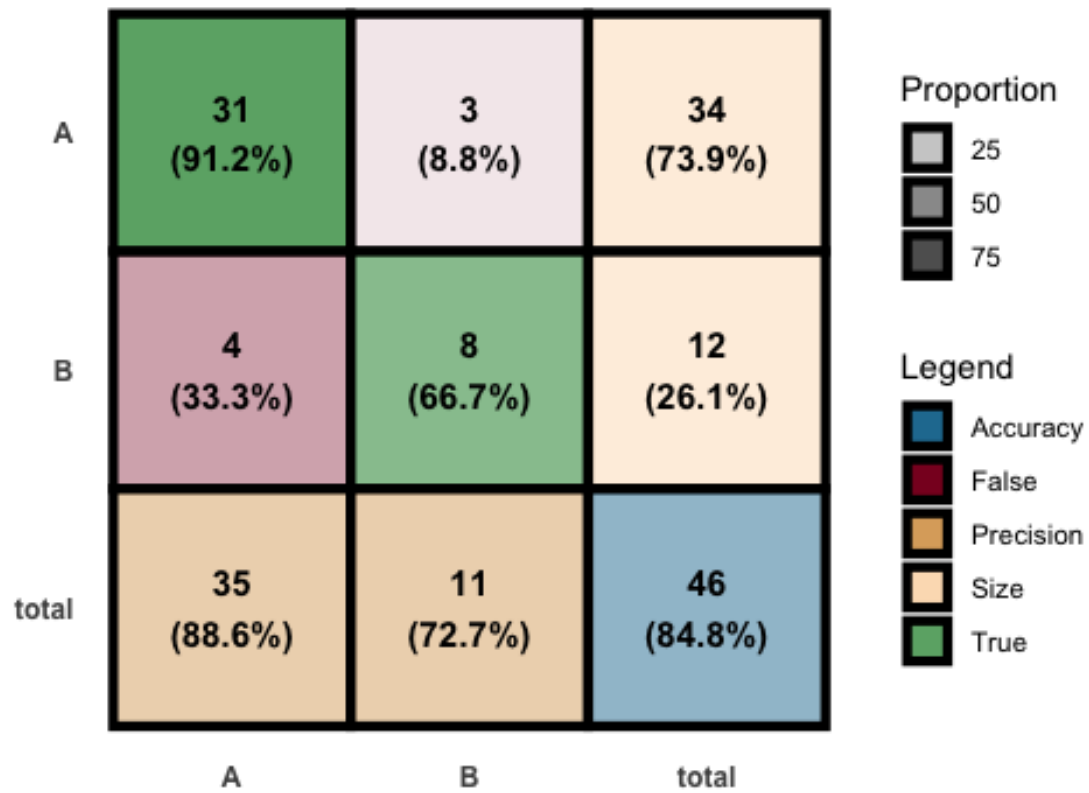
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
84.78%	0.518

Model Coefficients

	x
(Intercept)	1346.9848587
para_P	0.0650966
Dist(Hhal,C)_SM	-1333.2716154
fr_Ar_N1	-8.9378539
fr_Ar_N2	-46.9574822
fr_Ar_N3	-45.5774181
fr_amide1	-30.4305609
fr_amide2	-19.2214088

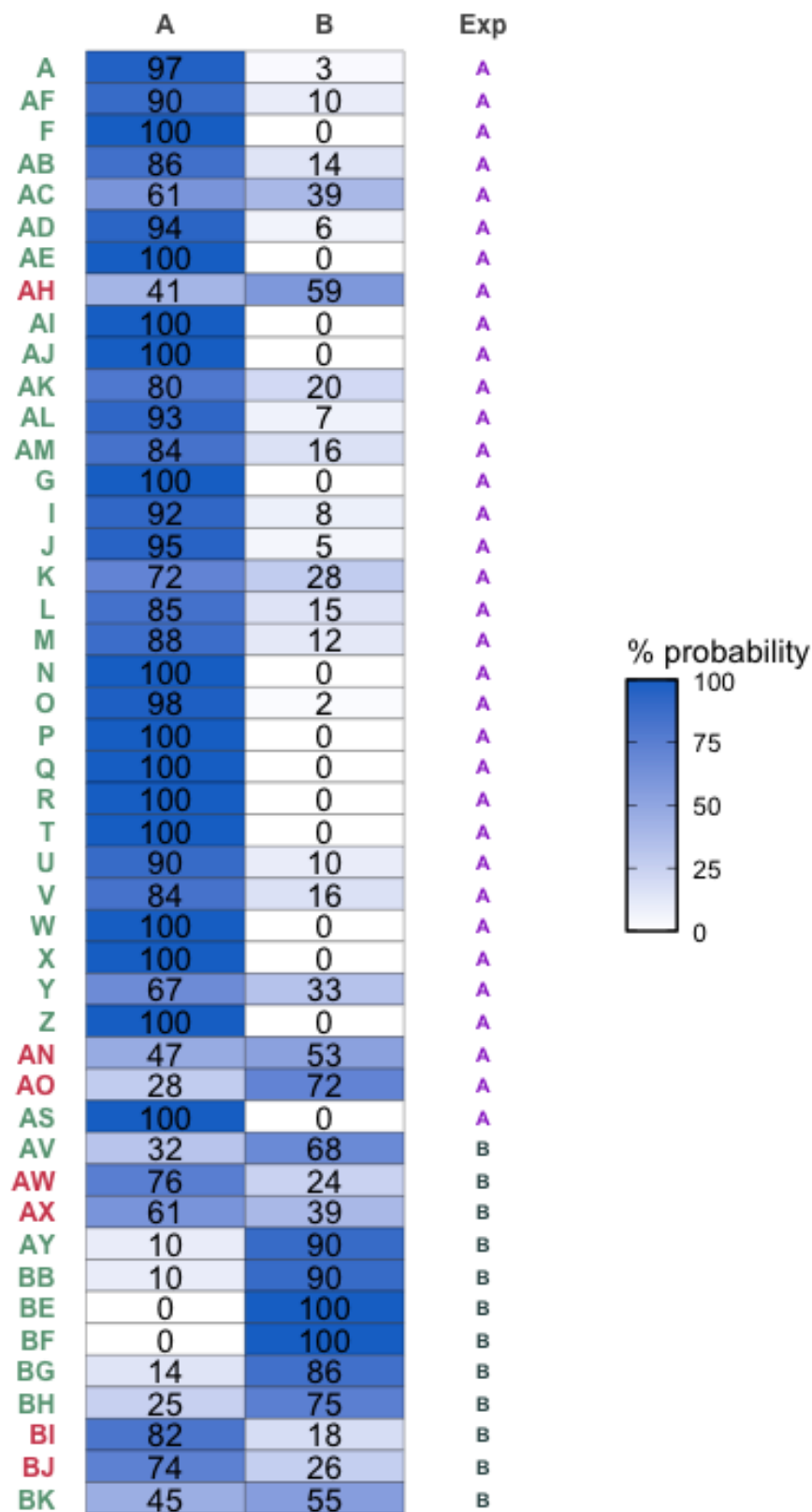
Training Set  
Confusion Matrix



3. 3rd Place

## Training Set

### Probability Heatmap



3. 3rd Place

## Test Set






### Confusion Matrix

A	<b>9</b> <b>(75%)</b>	<b>3</b> <b>(25%)</b>	<b>12</b> <b>(75%)</b>
B	<b>2</b> <b>(50%)</b>	<b>2</b> <b>(50%)</b>	<b>4</b> <b>(25%)</b>
total	<b>11</b> <b>(81.8%)</b>	<b>5</b> <b>(40%)</b>	<b>16</b> <b>(68.8%)</b>
	A	B	total

### Legend

	Accuracy
	False
	Precision
	Size
	True

### Proportion

	30
	40
	50
	60
	70

3. 3rd Place

## Test Set

### Probability Heatmap

	A	B
B	65	35
C	100	0
D	100	0
E	100	0
AA	13	87
H	76	24
S	100	0
AP	41	59
AQ	100	0
AR	43	57
AT	100	0
AU	60	40
AZ	38	62
BA	2	98
BC	93	7
BD	100	0

Exp

A

A

A

A

A

A

A

A

A

A

A

A

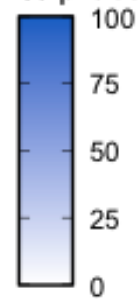
B

B

B

B

% probability



3. 3rd Place



#### 4th Place

Table:

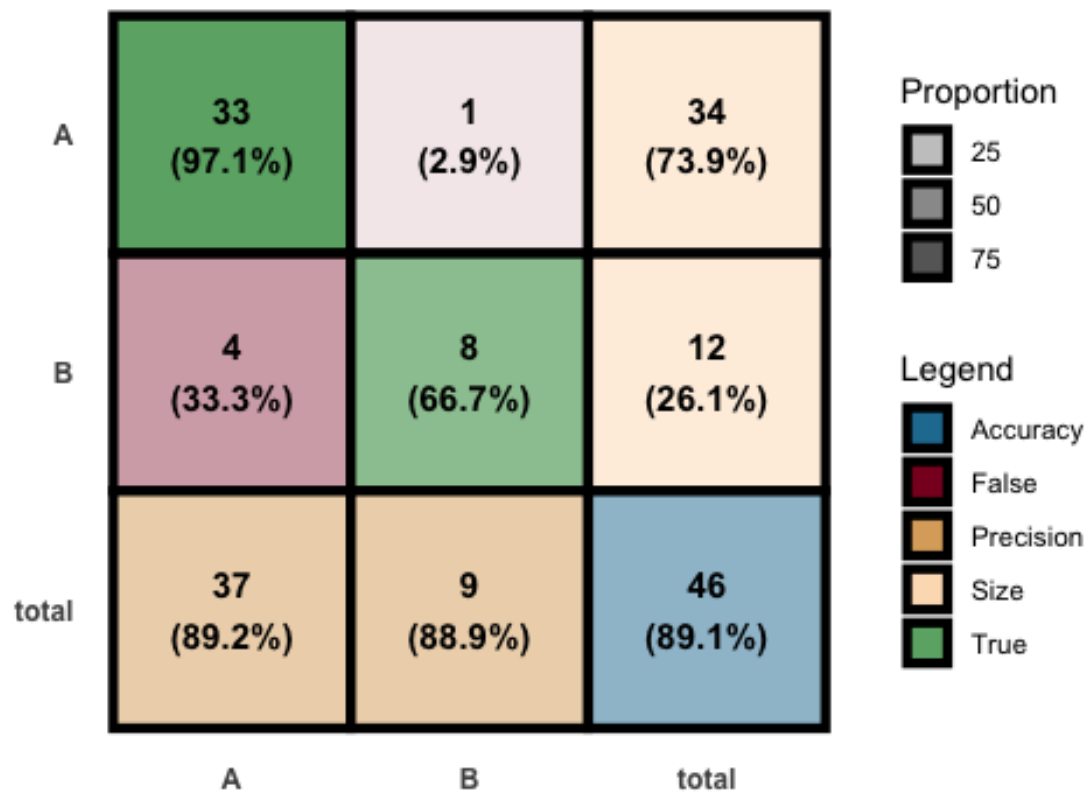
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
89.13%	0.512

Model Coefficients

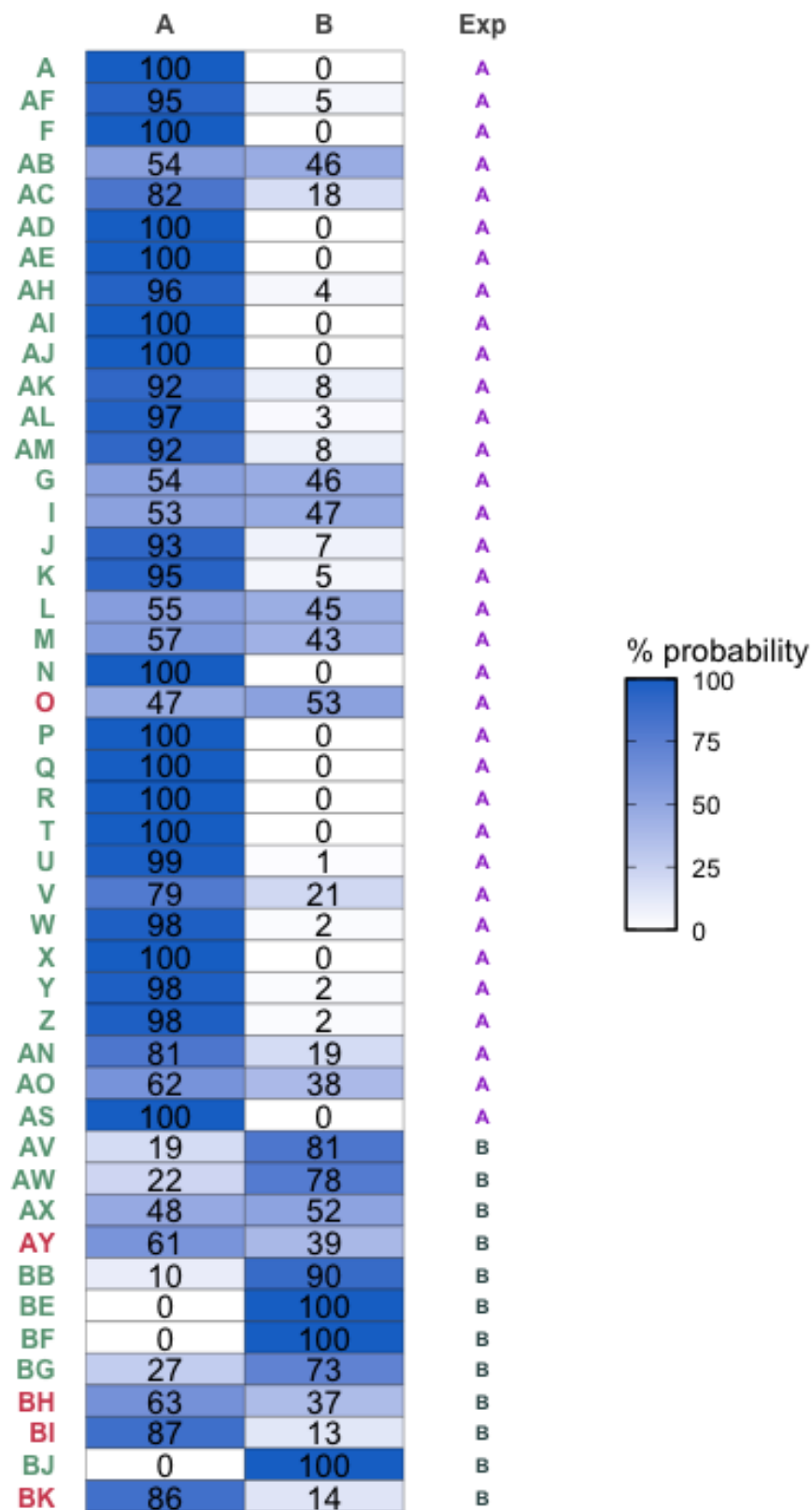
	x
(Intercept)	13.685359
NPA_sum_SM	8.233292
fr_Ar_N1	-6.823846
fr_Ar_N2	-19.015003
fr_Ar_N3	-18.765692
fr_aryl_methyl1	-3.779720
fr_aryl_methyl2	-11.798166
fr_benzene1	-1.329836
fr_benzene2	-15.621524

# Training Set Confusion Matrix



## Training Set

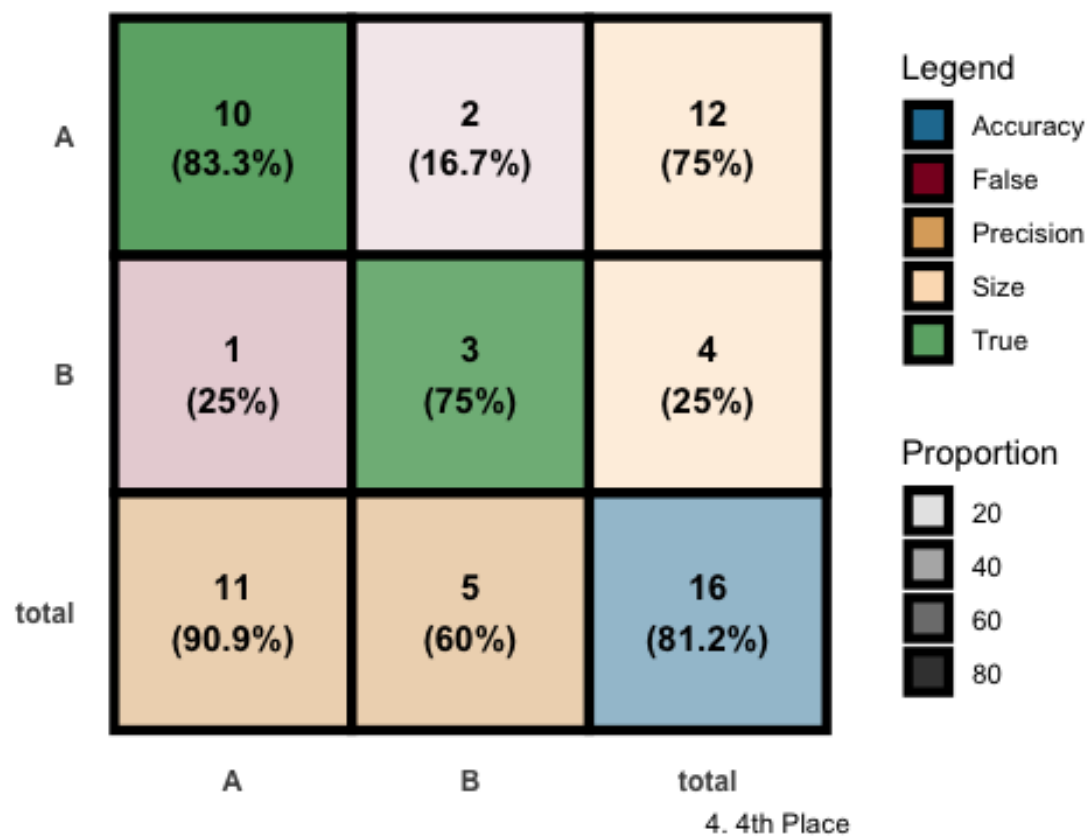
### Probability Heatmap



4. 4th Place

## Test Set

### Confusion Matrix



## Test Set

### Probability Heatmap

	A	B
<b>B</b>	46	54
<b>C</b>	100	0
<b>D</b>	100	0
<b>E</b>	100	0
<b>AA</b>	62	38
<b>H</b>	22	78
<b>S</b>	100	0
<b>AP</b>	84	16
<b>AQ</b>	85	15
<b>AR</b>	100	0
<b>AT</b>	100	0
<b>AU</b>	83	17
<b>AZ</b>	24	76
<b>BA</b>	0	100
<b>BC</b>	52	48
<b>BD</b>	2	98

Exp

A

A

A

A

A

A

A

A

A

A

A

A

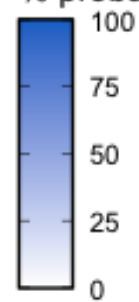
B

B

B

B

% probability



4. 4th Place

## 5th Place

Table:

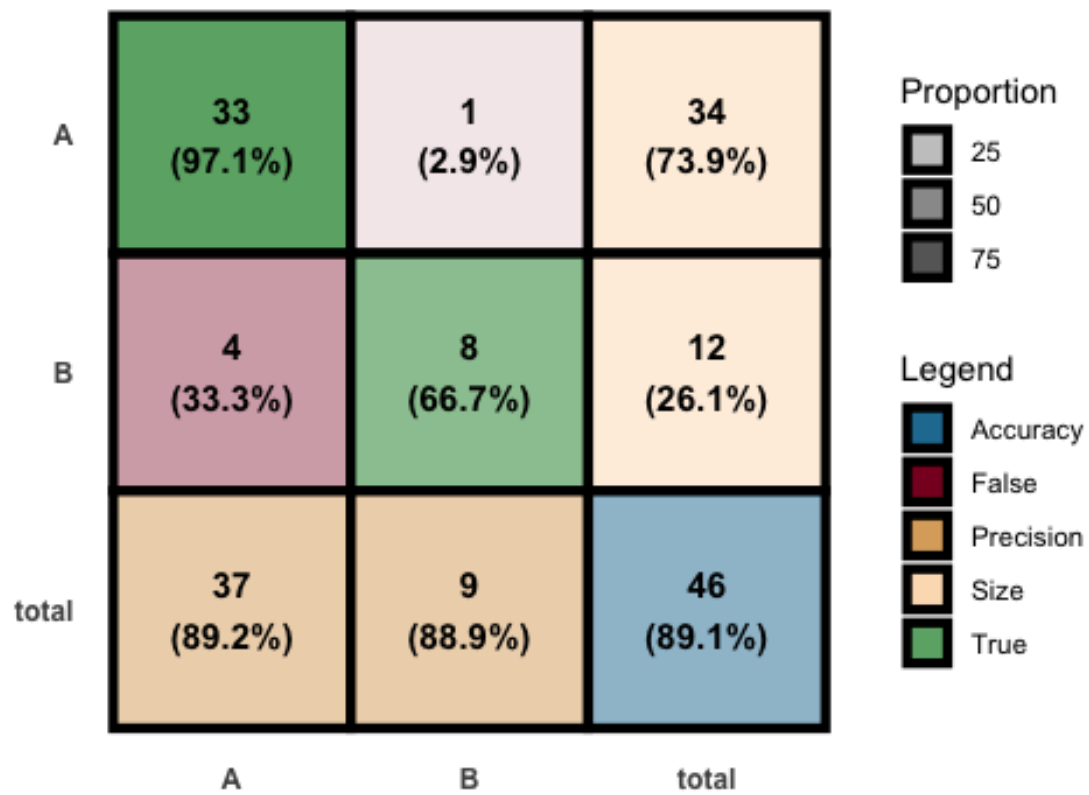
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
89.13%	0.512

Model Coefficients

	x
(Intercept)	43.957001
Avg_NPA_SM	49.387661
fr_Ar_N1	-37.097092
fr_Ar_N2	-69.144375
fr_Ar_N3	-69.553894
fr_aryl_methyl1	-3.779621
fr_aryl_methyl2	-51.951799
fr_benzene1	-1.330087
fr_benzene2	-68.569342

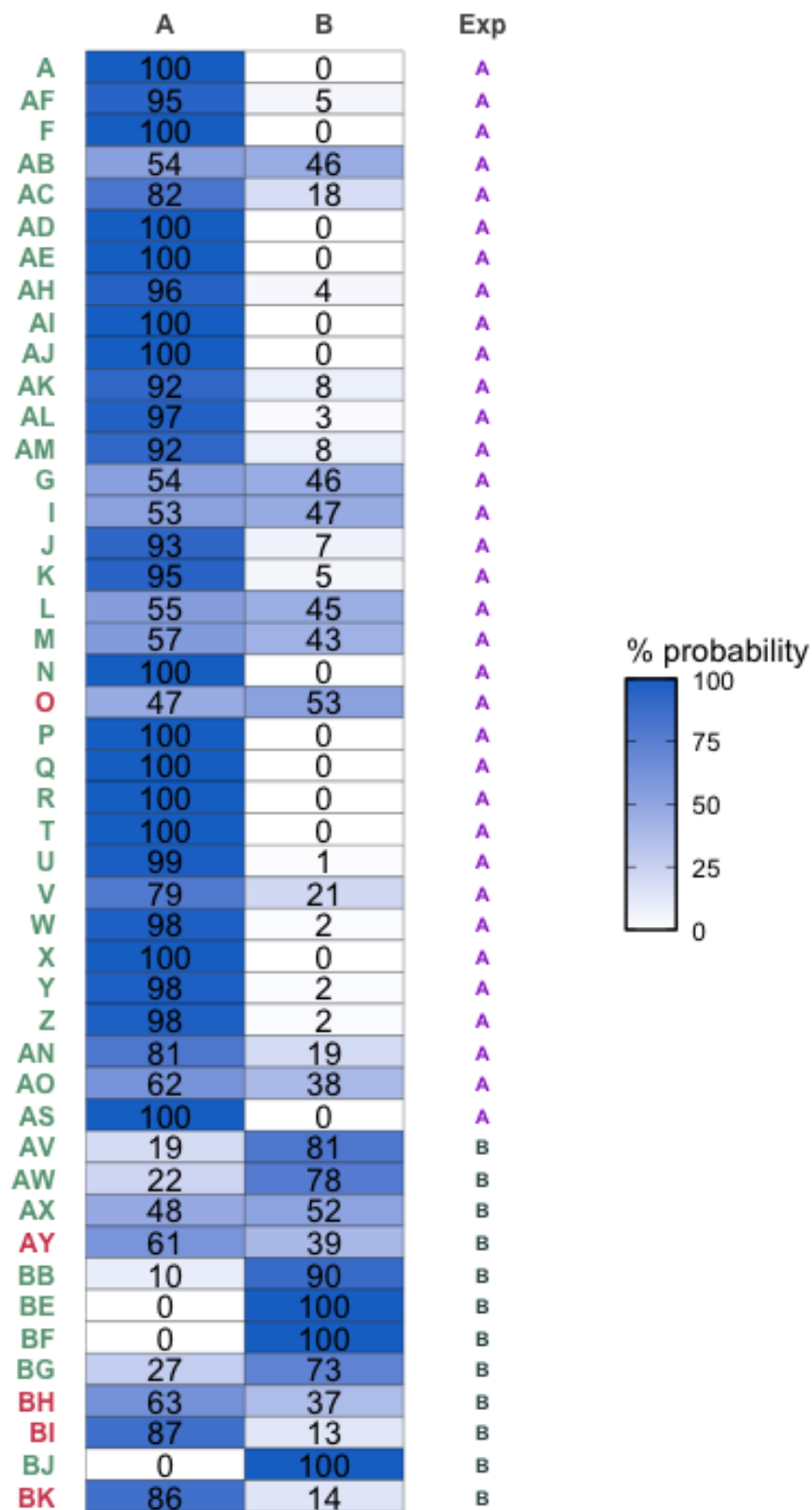
# Training Set Confusion Matrix



5. 5th Place

## Training Set

### Probability Heatmap

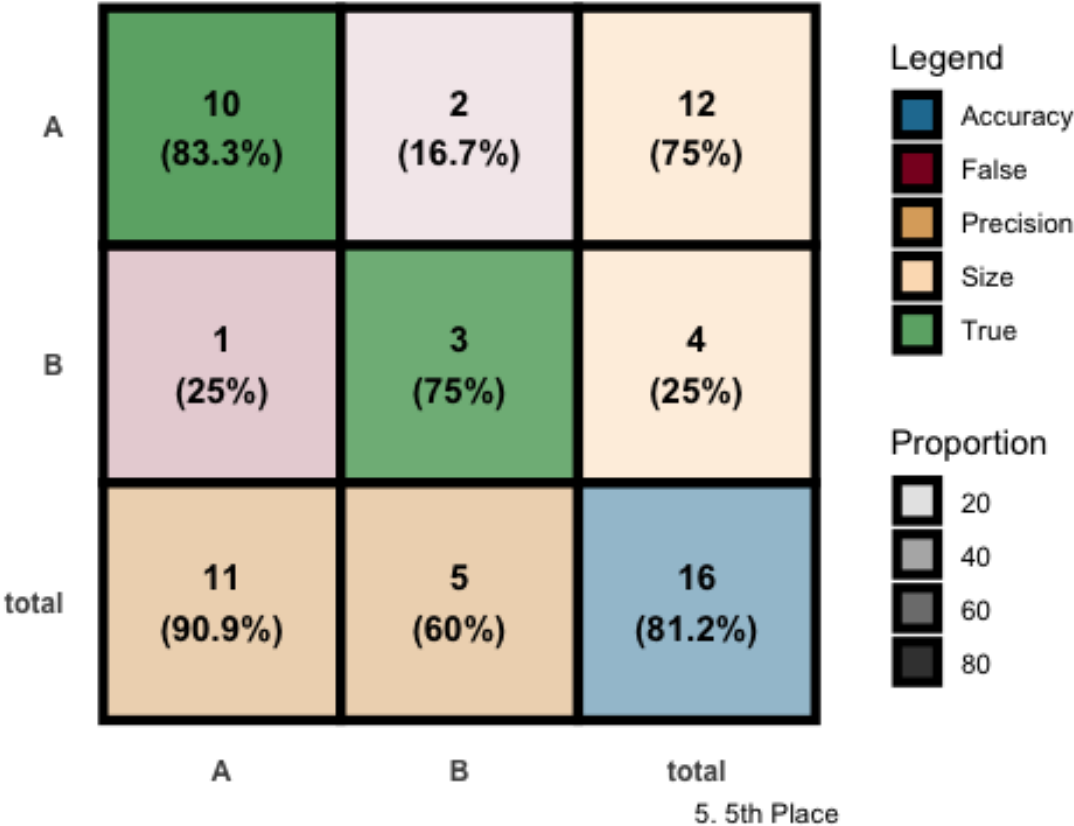


5. 5th Place



Test Set

Confusion Matrix



## Test Set

### Probability Heatmap

	A	B
B	46	54
C	100	0
D	100	0
E	100	0
AA	62	38
H	22	78
S	100	0
AP	84	16
AQ	85	15
AR	100	0
AT	100	0
AU	83	17
AZ	24	76
BA	0	100
BC	52	48
BD	2	98

Exp

A

A

A

A

A

A

A

A

A

A

A

A

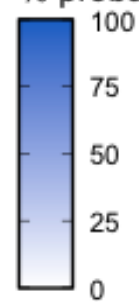
B

B

B

B

% probability



5. 5th Place

### 5th Place with 5 features (fr\_halogen added)

Table:

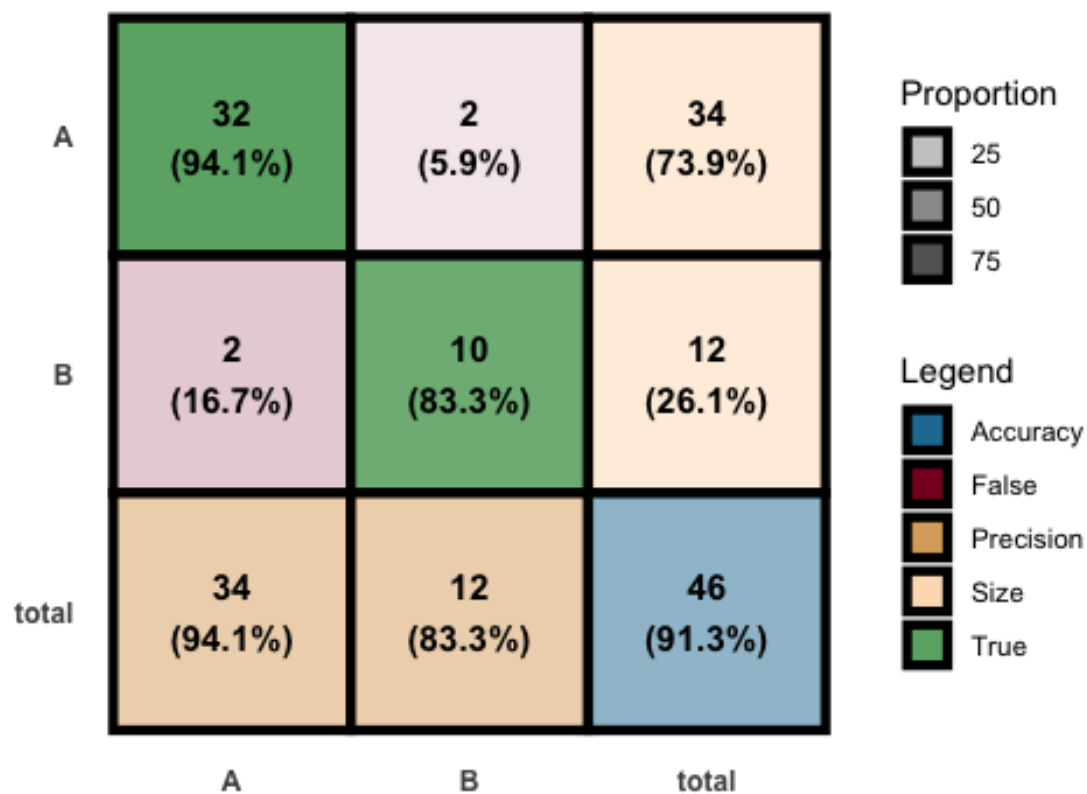
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
91.3%	0.578

Model Coefficients

	x
(Intercept)	38.099048
Avg_NPA_SM	48.412036
fr_Ar_N1	-31.903705
fr_Ar_N2	-60.673746
fr_Ar_N3	-66.149054
fr_aryl_methyl1	-4.592317
fr_aryl_methyl2	-37.006653
fr_benzene1	-1.970290
fr_benzene2	-77.887220
fr_halogen1	1.607589
fr_halogen2	2.929693
fr_halogen3	2.595277
fr_halogen4	-2.087011
fr_halogen5	2.187832
fr_halogen6	-2.311938

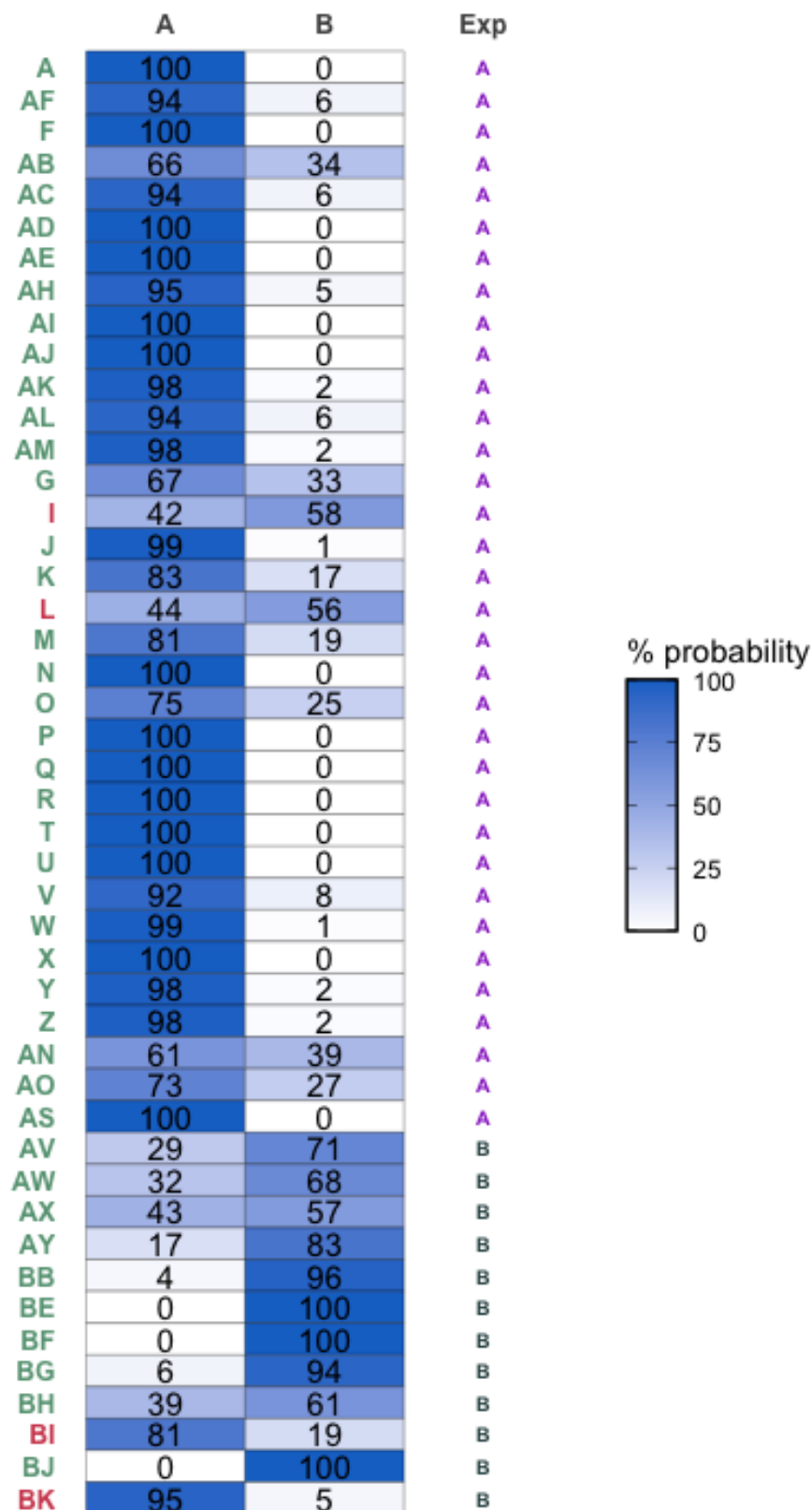
# Training Set Confusion Matrix



5. 5th Place - 5 features

## Training Set

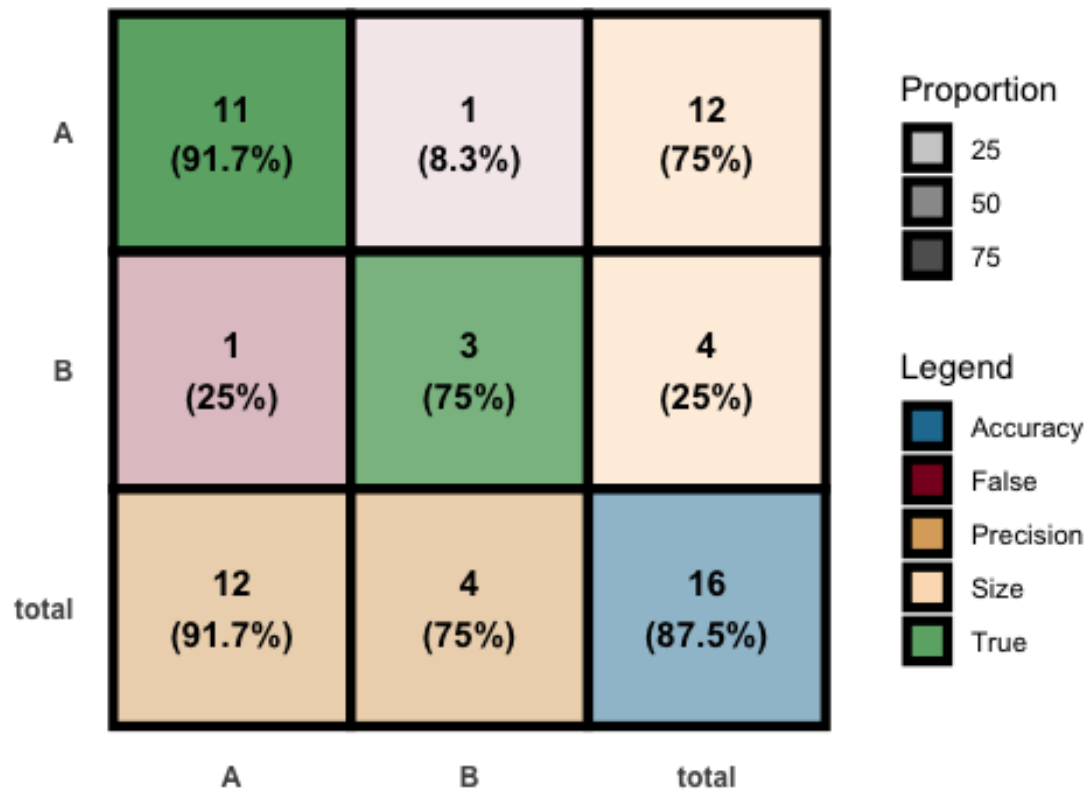
### Probability Heatmap



5. 5th Place - 5 features

## Test Set

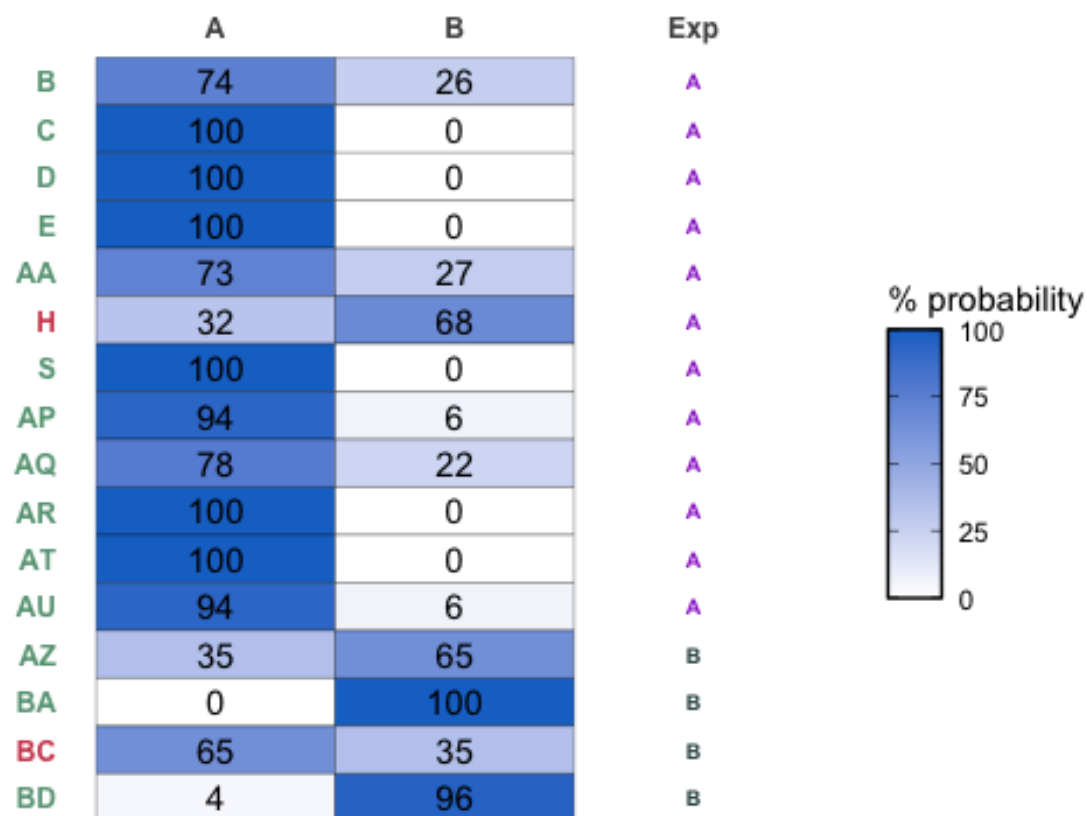
### Confusion Matrix



5. 5th Place - 5 features

## Test Set

### Probability Heatmap



5. 5th Place - 5 features

## 6th Place

Table:

Full Model Stats - Overall Accuracy and Pseudo-R2

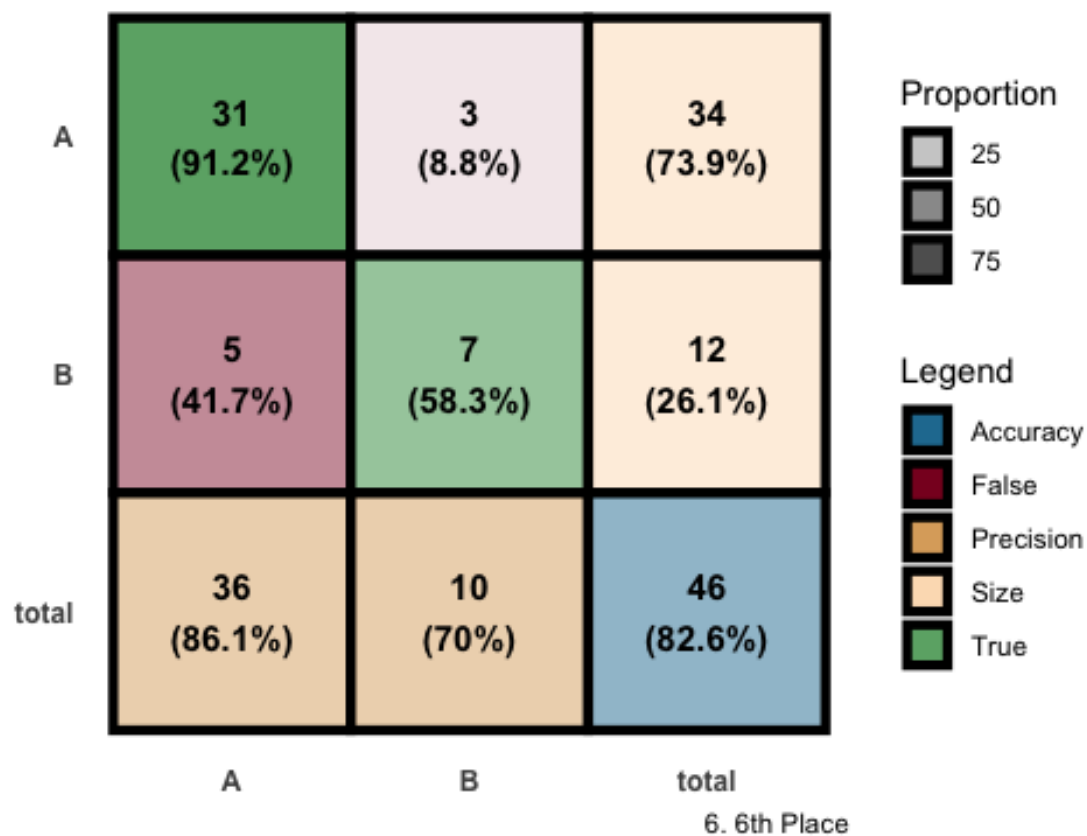
Accuracy	McFadden_R2
82.61%	0.509

Model Coefficients

	x
(Intercept)	18.352407
NPA_5_P	-37.318779
fr_Ar_N1	-30.290325
fr_Ar_N2	-51.893934
fr_Ar_N3	-52.664476
fr_aryl_methyl1	-3.184576
fr_aryl_methyl2	-38.499319
fr_benzene1	-2.336220
fr_benzene2	-50.576774

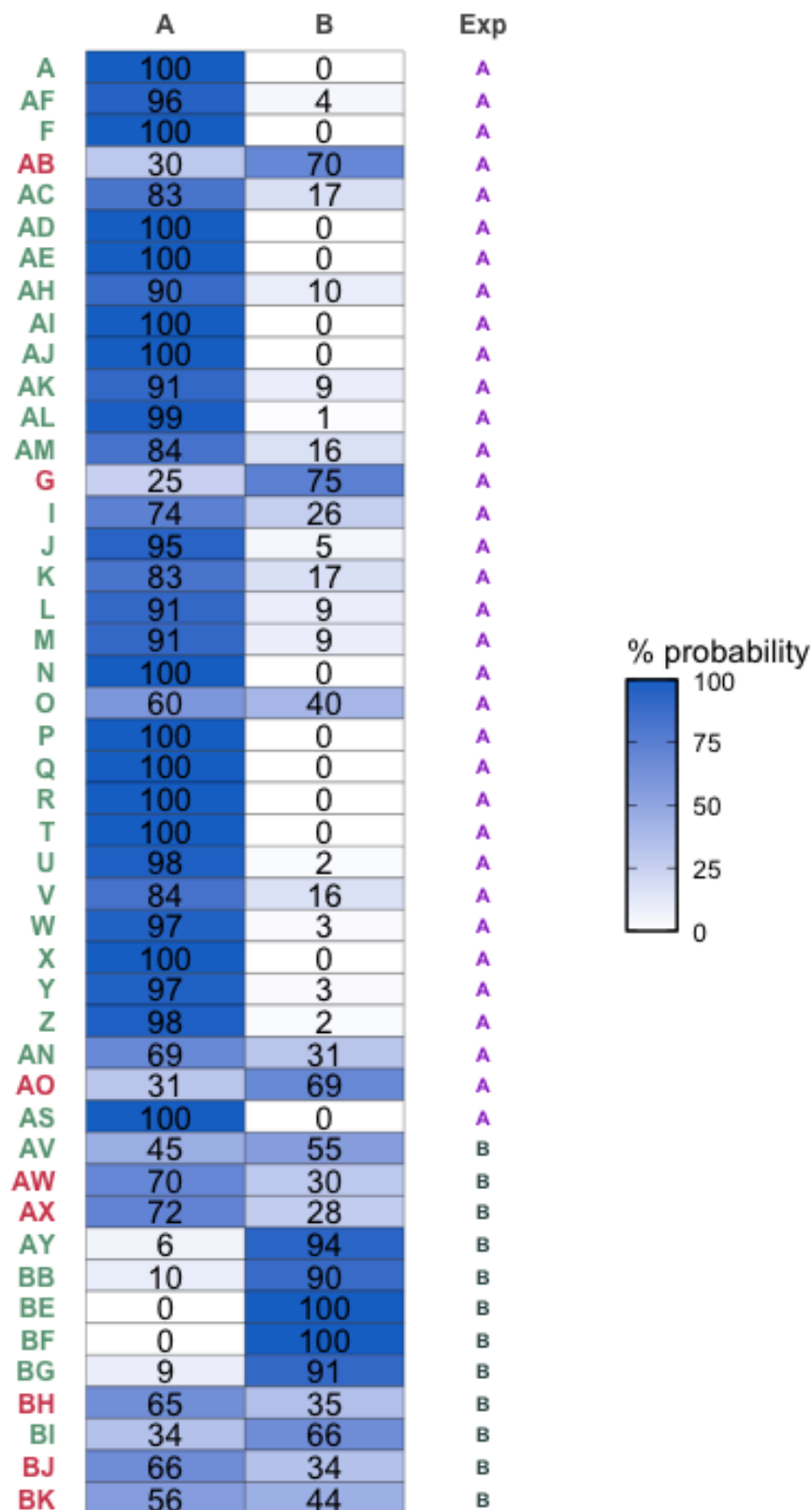


# Training Set Confusion Matrix



# Training Set

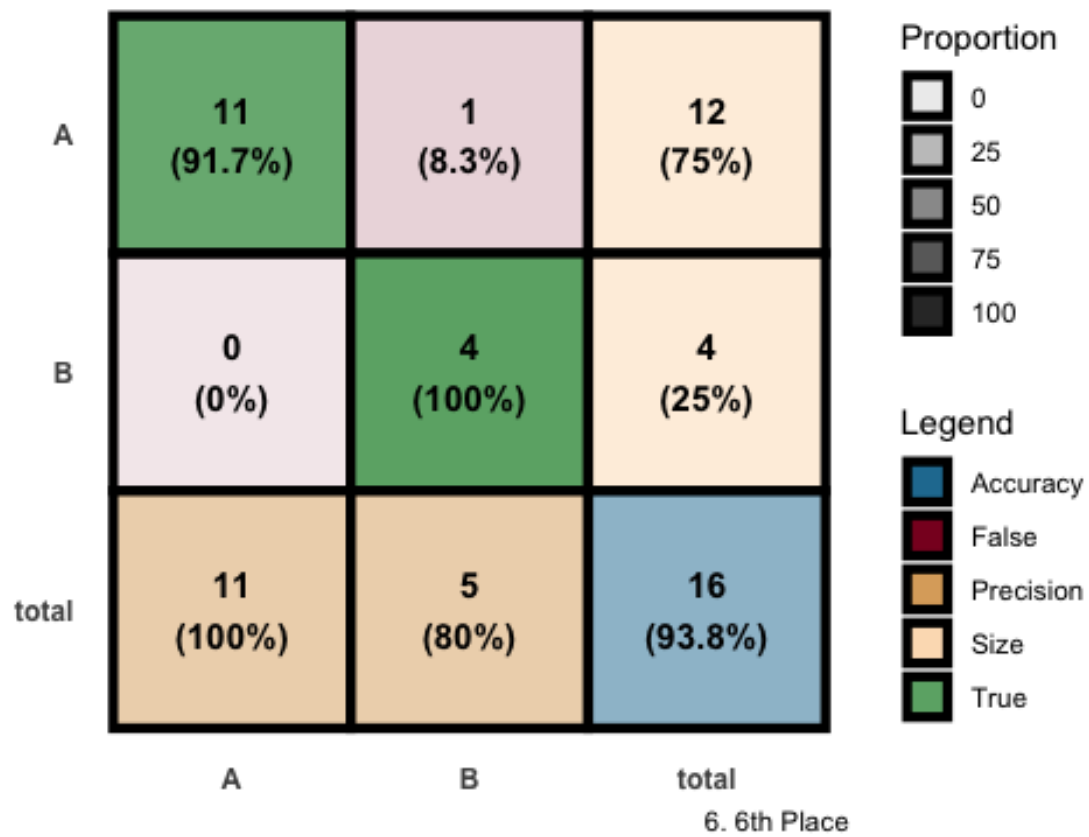
## Probability Heatmap



6. 6th Place

## Test Set

### Confusion Matrix



## Test Set

### Probability Heatmap

	A	B
B	86	14
C	100	0
D	100	0
E	100	0
AA	25	75
H	70	30
S	100	0
AP	57	43
AQ	59	41
AR	100	0
AT	100	0
AU	67	33
AZ	1	99
BA	45	55
BC	35	65
BD	12	88

Exp

A

A

A

A

A

A

A

A

A

A

A

A

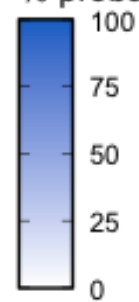
B

B

B

B

% probability



6. 6th Place

## 7th Place

Table:

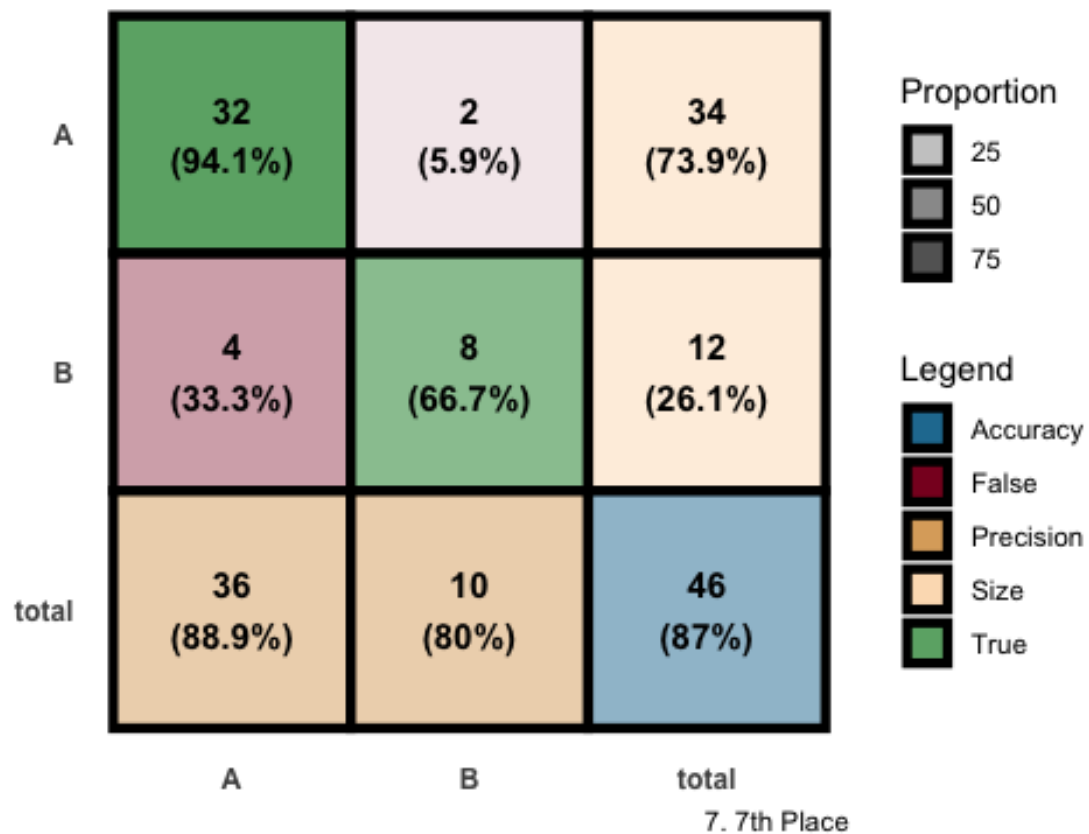
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
86.96%	0.504

Model Coefficients

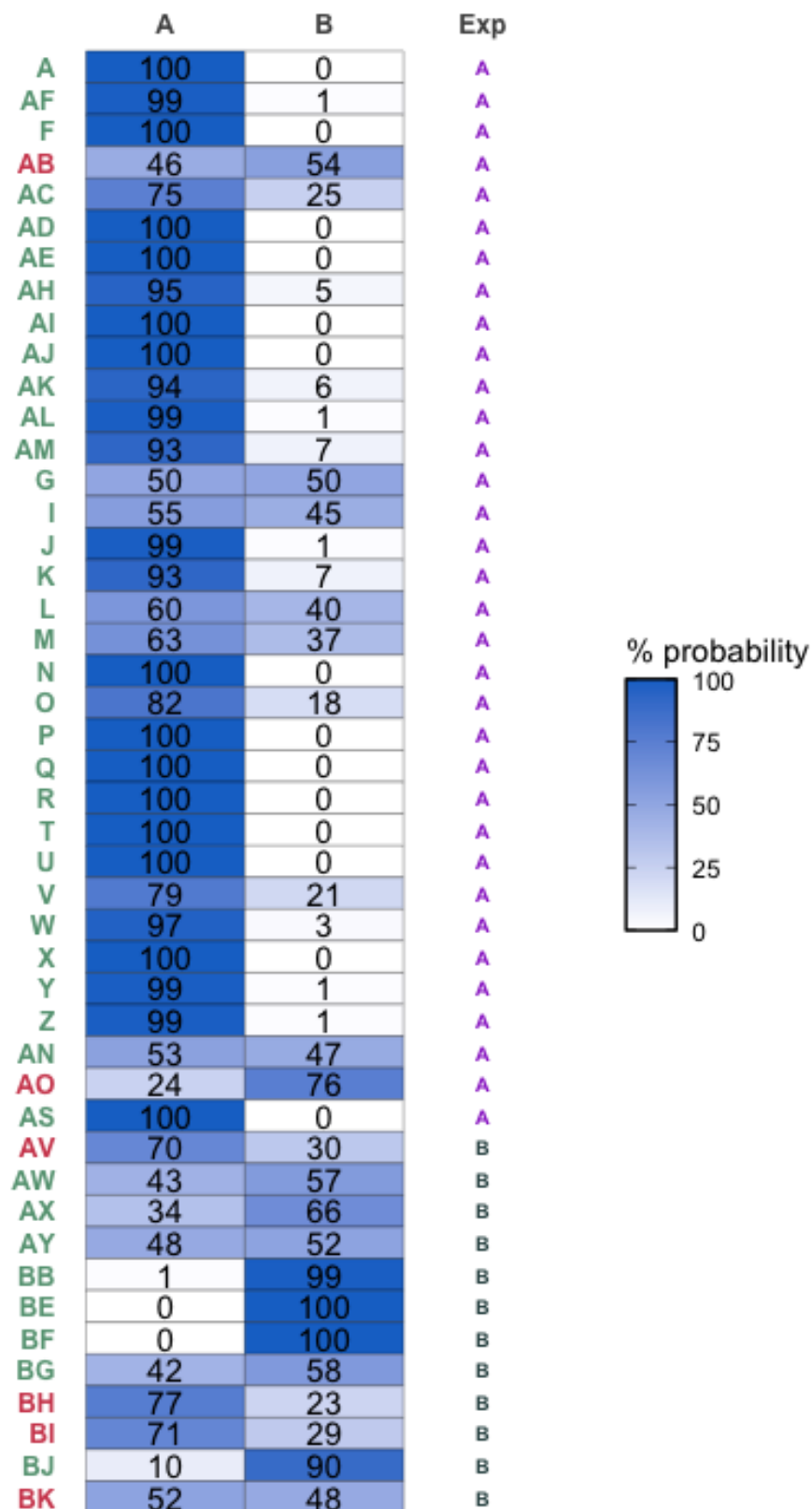
	x
(Intercept)	8.287436
NPA_6_SM	9.063678
fr_Ar_N1	-8.593052
fr_Ar_N2	-18.422215
fr_Ar_N3	-17.239187
fr_aryl_methyl1	-4.379548
fr_aryl_methyl2	-9.520873
fr_benzene1	-1.447618
fr_benzene2	-12.085137

# Training Set Confusion Matrix



# Training Set

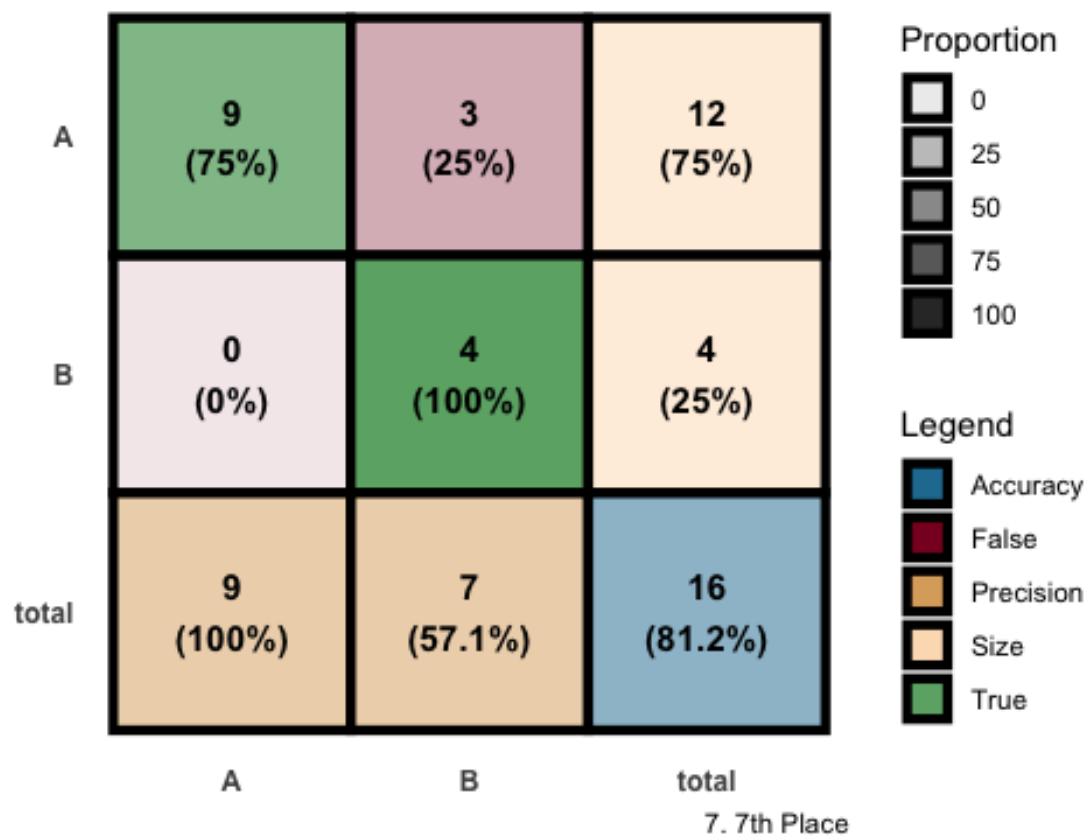
## Probability Heatmap



7. 7th Place

## Test Set

### Confusion Matrix





## Test Set

### Probability Heatmap

	A	B
B	76	24
C	100	0
D	100	0
E	100	0
AA	40	60
H	43	57
S	100	0
AP	47	53
AQ	80	20
AR	100	0
AT	100	0
AU	52	48
AZ	1	99
BA	47	53
BC	49	51
BD	8	92

Exp

A

A

A

A

A

A

A

A

A

A

A

A

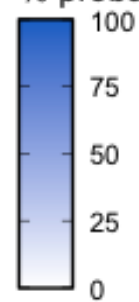
B

B

B

B

% probability



7. 7th Place

## 8th Place

Table:

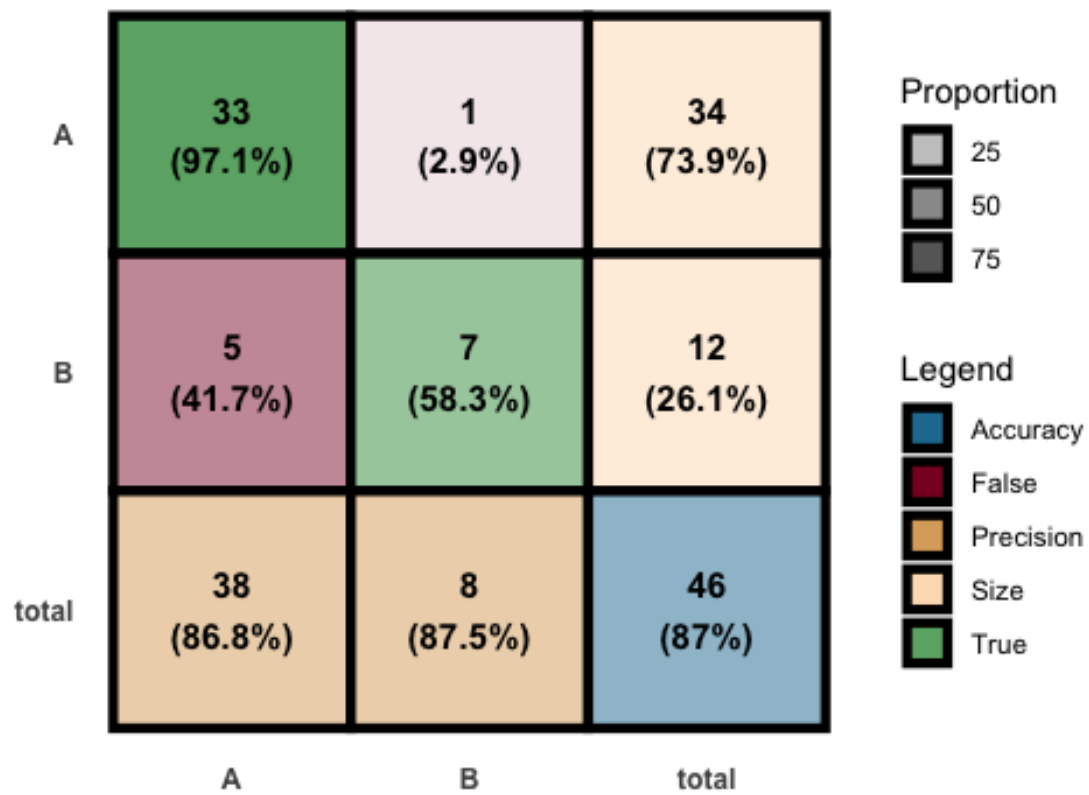
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
86.96%	0.49

Model Coefficients

	x
(Intercept)	1116.704277
Avg_NPA_SM	40.417204
Dist(Hha1,C)_SM	-1020.317722
fr_Ar_N1	-7.001472
fr_Ar_N2	-33.667037
fr_Ar_N3	-32.106211
fr_aryl_methyl1	-3.048454
fr_aryl_methyl2	-20.561621

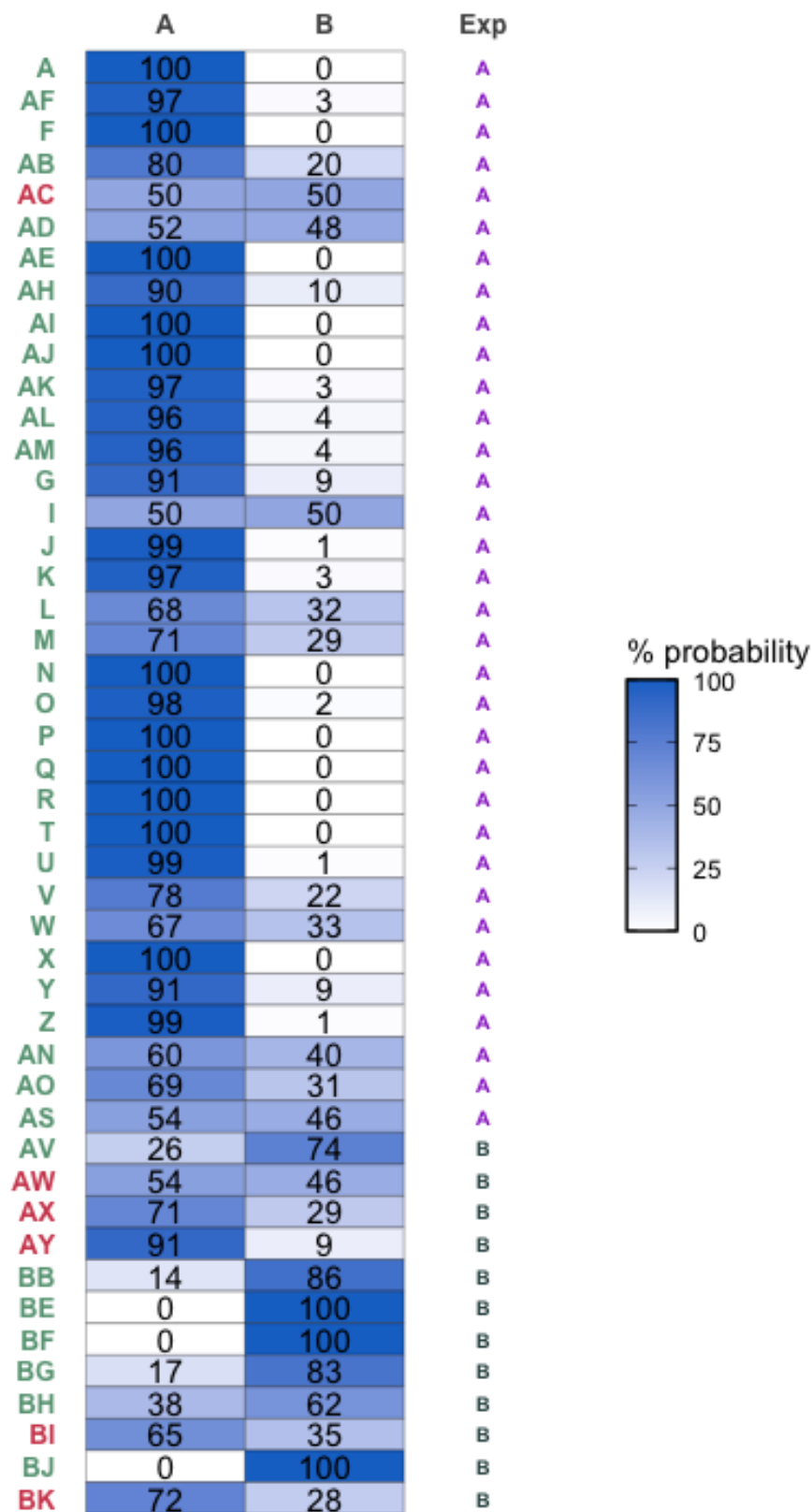
# Training Set Confusion Matrix



8. 8th Place

# Training Set

## Probability Heatmap



8. 8th Place

## Test Set

### Confusion Matrix

A	<b>10</b> <b>(83.3%)</b>	<b>2</b> <b>(16.7%)</b>	<b>12</b> <b>(75%)</b>
B	<b>1</b> <b>(25%)</b>	<b>3</b> <b>(75%)</b>	<b>4</b> <b>(25%)</b>
total	<b>11</b> <b>(90.9%)</b>	<b>5</b> <b>(60%)</b>	<b>16</b> <b>(81.2%)</b>
	A	B	total

### Legend

	Accuracy
	False
	Precision
	Size
	True

### Proportion

	20
	40
	60
	80

8. 8th Place

## Test Set

### Probability Heatmap

	A	B
B	31	69
C	100	0
D	100	0
E	100	0
AA	47	53
H	54	46
S	100	0
AP	71	29
AQ	75	25
AR	64	36
AT	100	0
AU	66	34
AZ	37	63
BA	0	100
BC	92	8
BD	4	96

Exp

A

A

A

A

A

A

A

A

A

A

A

A

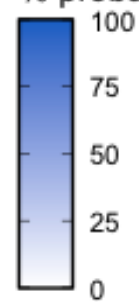
B

B

B

B

% probability



8. 8th Place

## 9th Place

Table:

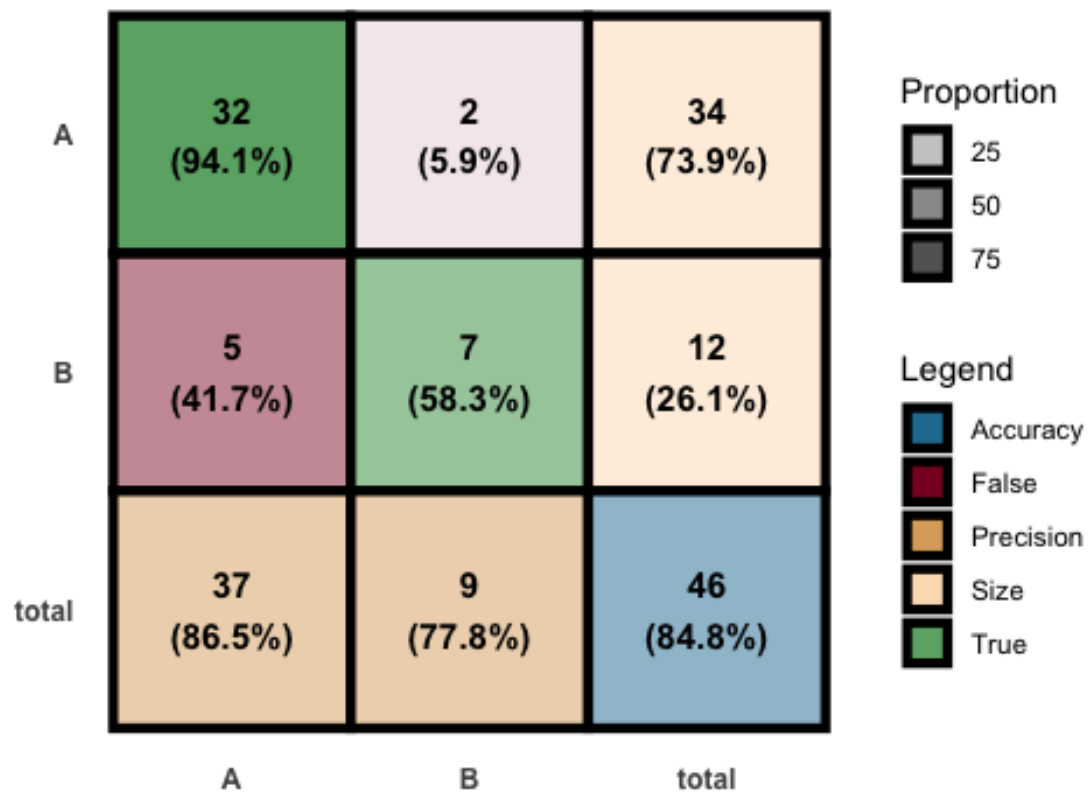
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
84.78%	0.483

Model Coefficients

	x
(Intercept)	6.9649782
para.angle_P	-0.0752978
NPA_6_SM	9.5322216
fr_Ar_N1	-6.7180680
fr_Ar_N2	-14.6849449
fr_Ar_N3	-10.6206149
fr_aryl_methyl1	-3.8589971
fr_aryl_methyl2	-7.5751954

Training Set  
Confusion Matrix

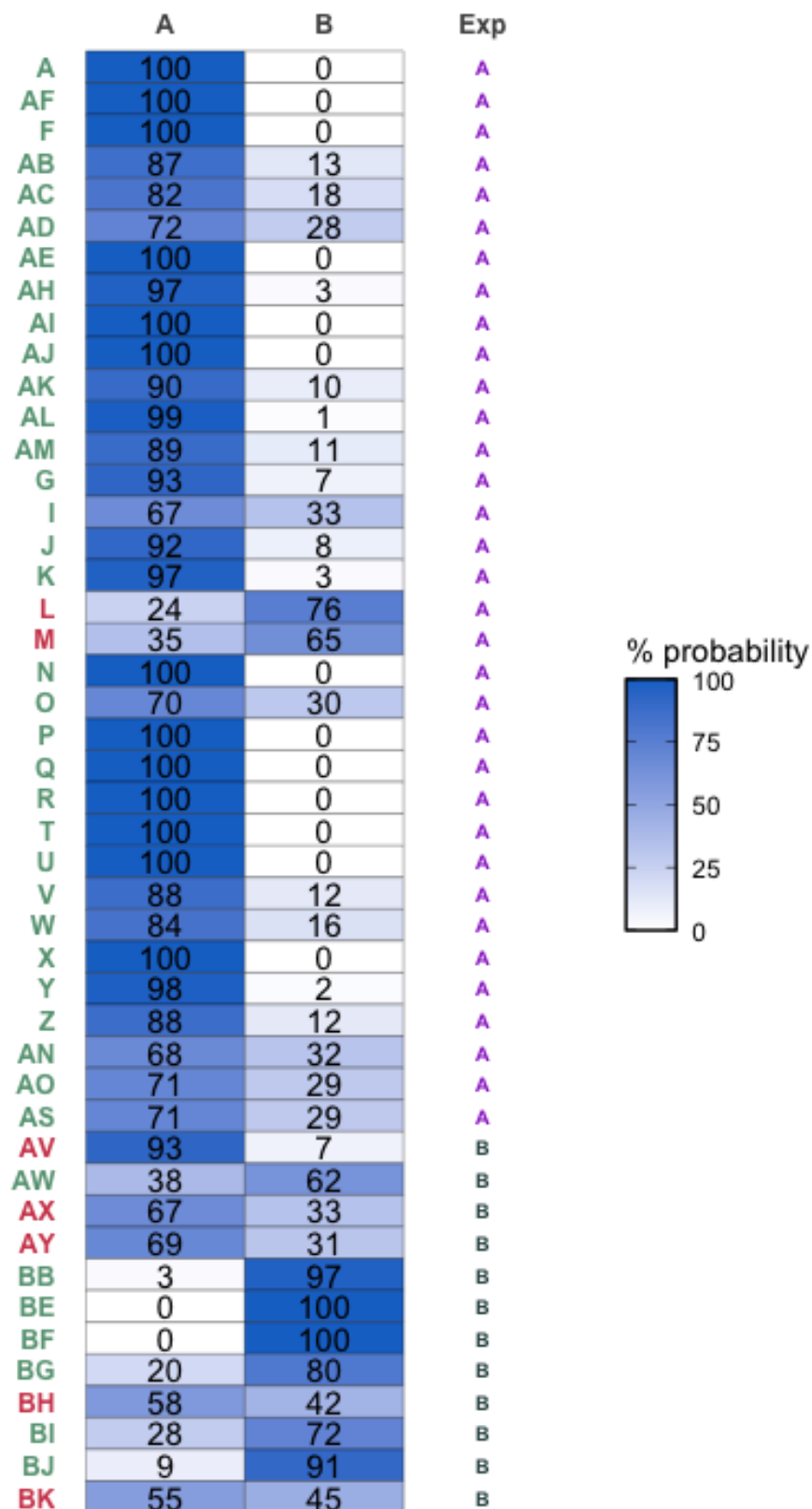


9. 9th Place



# Training Set

## Probability Heatmap




9. 9th Place

## Test Set





### Confusion Matrix

A	<b>10</b> (83.3%)	<b>2</b> (16.7%)	<b>12</b> (75%)
B	<b>2</b> (50%)	<b>2</b> (50%)	<b>4</b> (25%)
total	<b>12</b> (83.3%)	<b>4</b> (50%)	<b>16</b> (75%)
	A	B	total

### Legend

	Accuracy
	False
	Precision
	Size
	True

### Proportion

	20
	40
	60
	80

9. 9th Place

## Test Set

### Probability Heatmap

	A	B
B	63	37
C	100	0
D	100	0
E	100	0
AA	83	17
H	37	63
S	84	16
AP	59	41
AQ	91	9
AR	50	50
AT	99	1
AU	64	36
AZ	13	87
BA	92	8
BC	86	14
BD	34	66

Exp

A

A

A

A

A

A

A

A

A

A

A

A

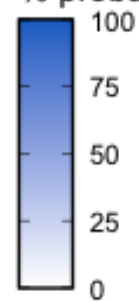
B

B

B

B

% probability



9. 9th Place

## 10th Place

Table:

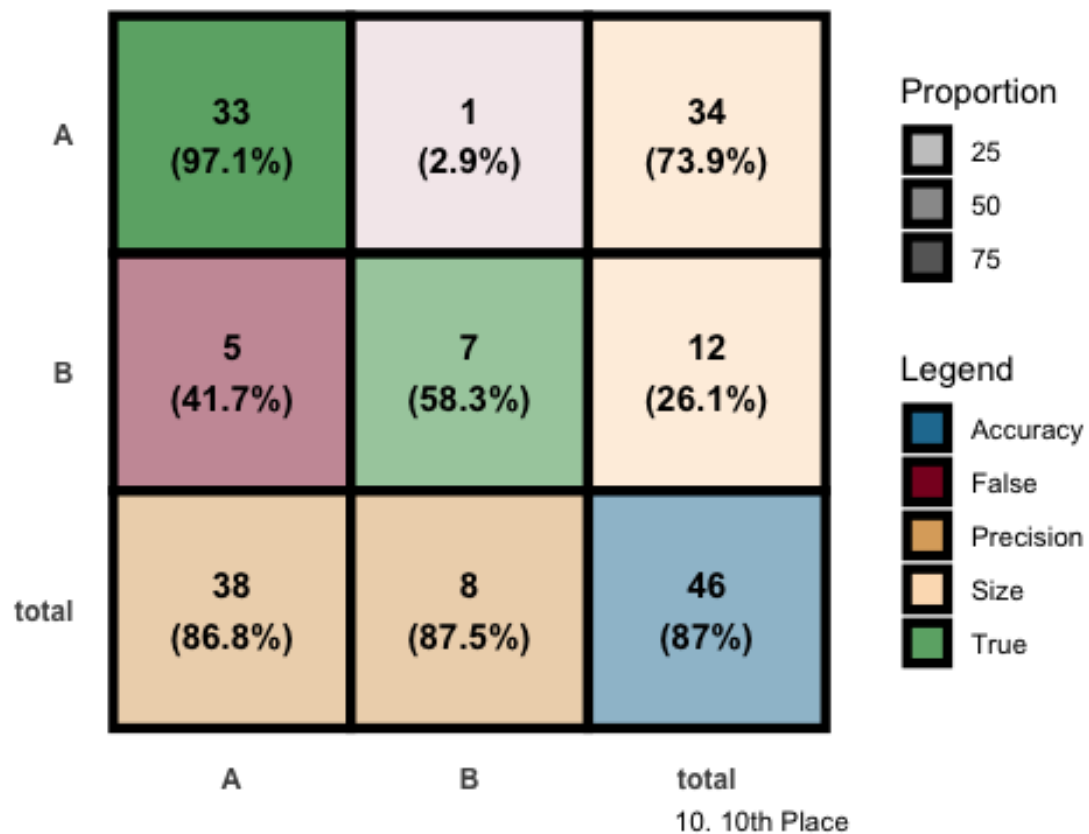
Full Model Stats - Overall Accuracy and Pseudo-R2

Accuracy	McFadden_R2
86.96%	0.483

Model Coefficients

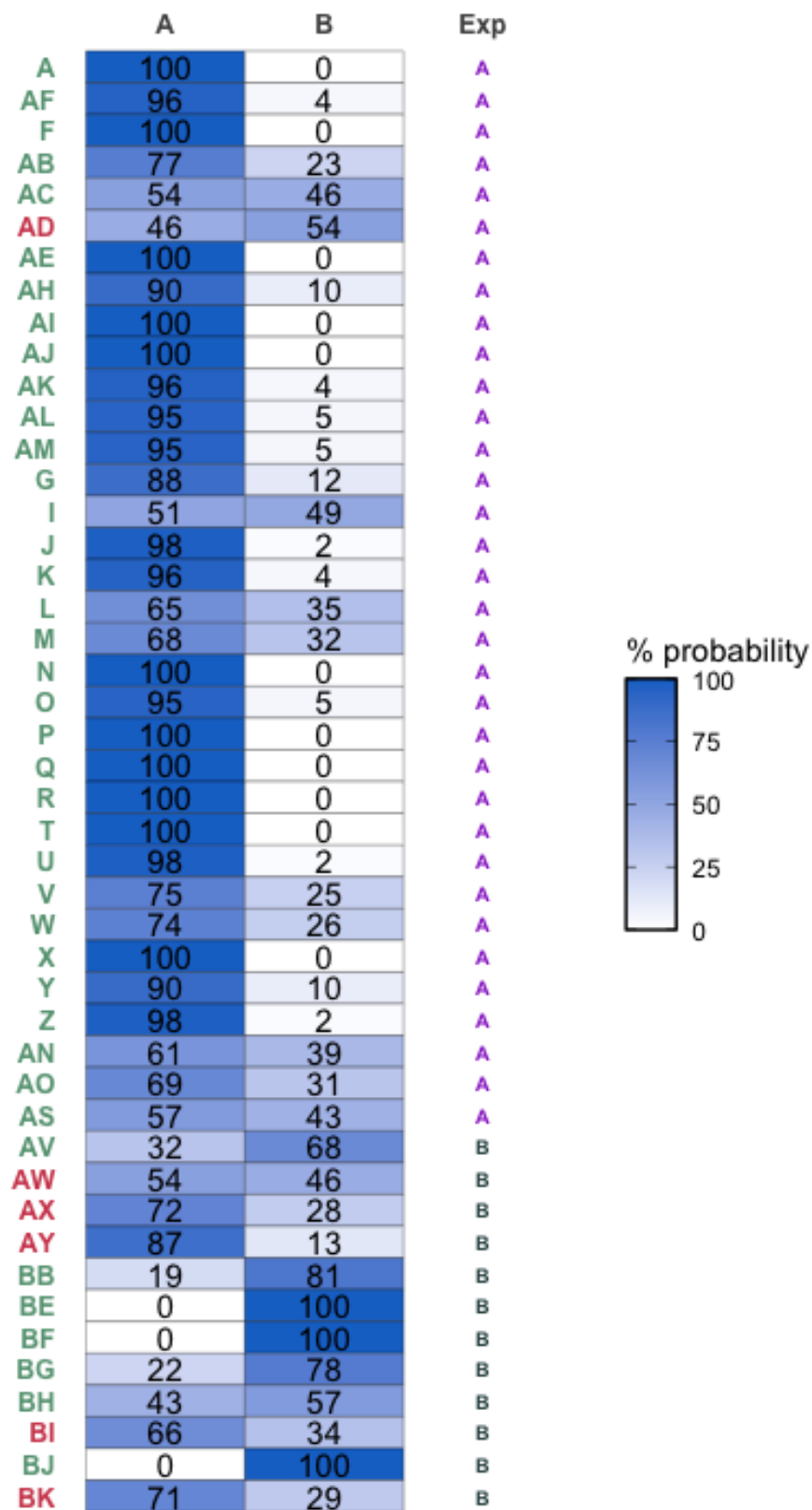
	x
(Intercept)	850.026460
NPA_sum_SM	5.827134
Dist(Hha1,C)_SM	-773.766151
fr_Ar_N1	-8.017928
fr_Ar_N2	-36.874714
fr_Ar_N3	-35.321092
fr_aryl_methyl1	-2.778137
fr_aryl_methyl2	-12.713464

# Training Set Confusion Matrix



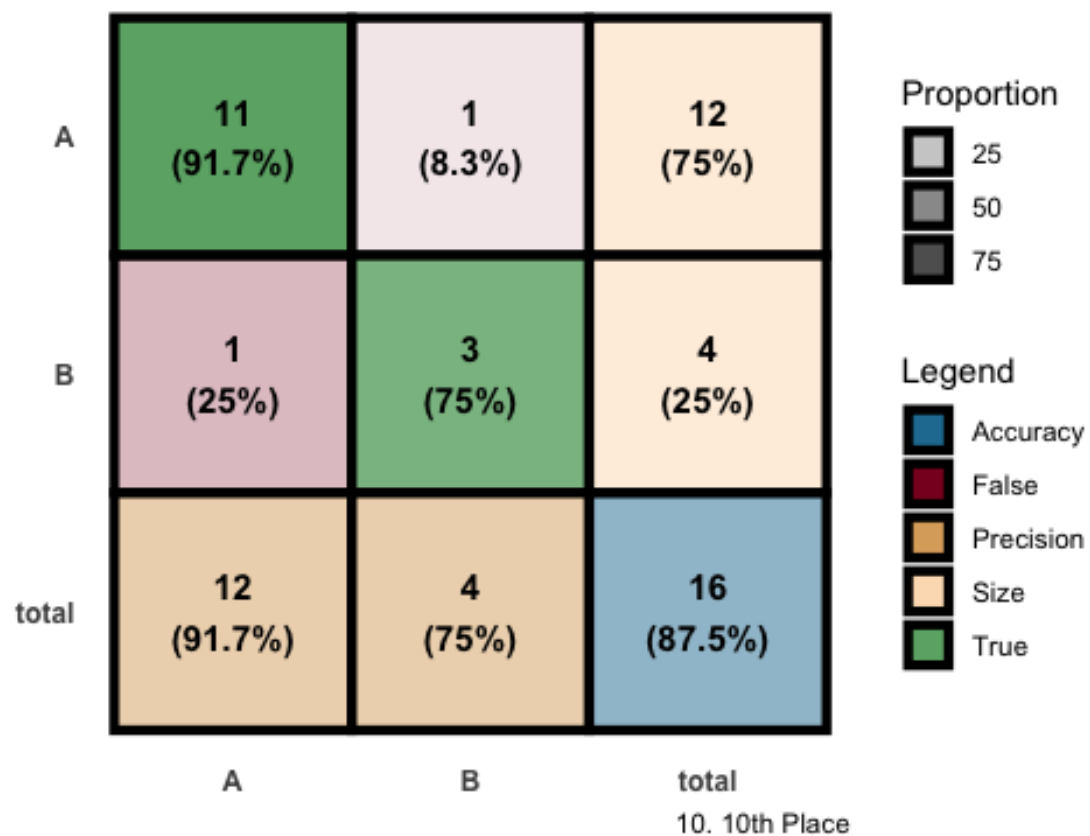
# Training Set

## Probability Heatmap



## Test Set

### Confusion Matrix



## Test Set

### Probability Heatmap

	A	B
<b>B</b>	35	65
<b>C</b>	100	0
<b>D</b>	100	0
<b>E</b>	100	0
<b>AA</b>	52	48
<b>H</b>	54	46
<b>S</b>	100	0
<b>AP</b>	70	30
<b>AQ</b>	74	26
<b>AR</b>	64	36
<b>AT</b>	100	0
<b>AU</b>	66	34
<b>AZ</b>	41	59
<b>BA</b>	0	100
<b>BC</b>	88	12
<b>BD</b>	6	94

Exp

A

A

A

A

A

A

A

A

A

A

A

A

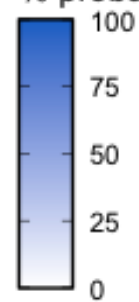
B

B

B

B

% probability



10. 10th Place