Aldehyde Deuteration - Best Model Feature Analysis

Table of Contents

[Model Search 1](#_Toc169612915)

[1st Place 1](#_Toc169612916)

[Model Analysis - Variable Impact 8](#_Toc169612917)

[1. NPA\_1 Removed 8](#_Toc169612918)

[2. Dist(2, 7) Removed 15](#_Toc169612919)

[3. dip\_y Removed 22](#_Toc169612920)

[4. -2-3- Removed 29](#_Toc169612921)

### Model Search

### 1st Place

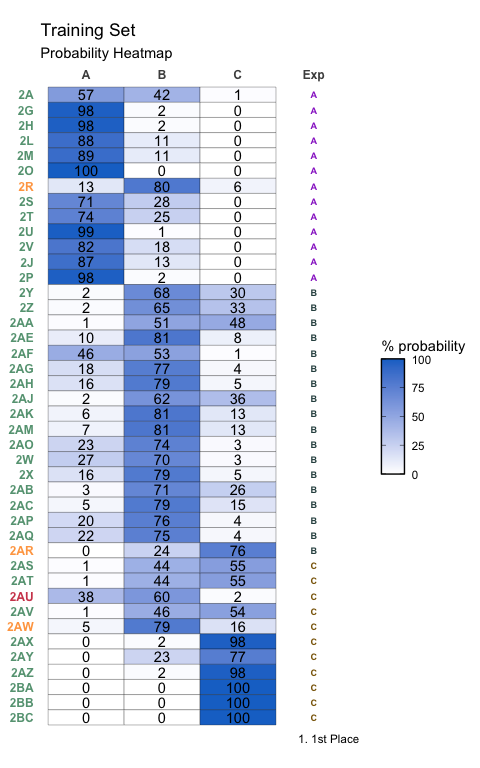
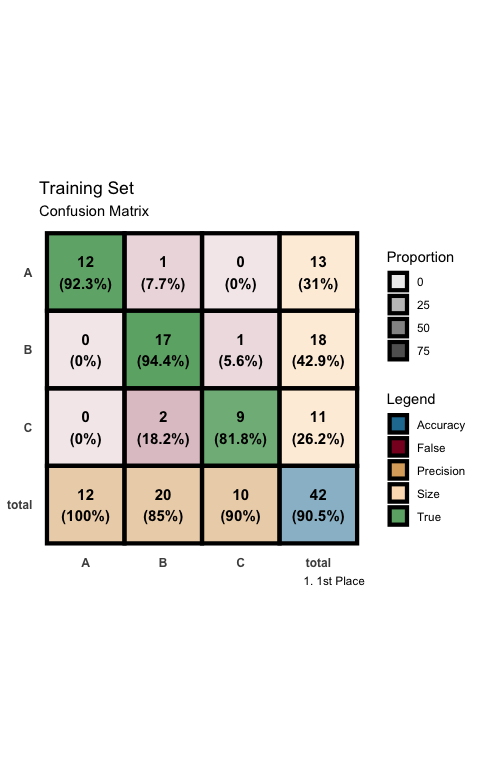
Table:

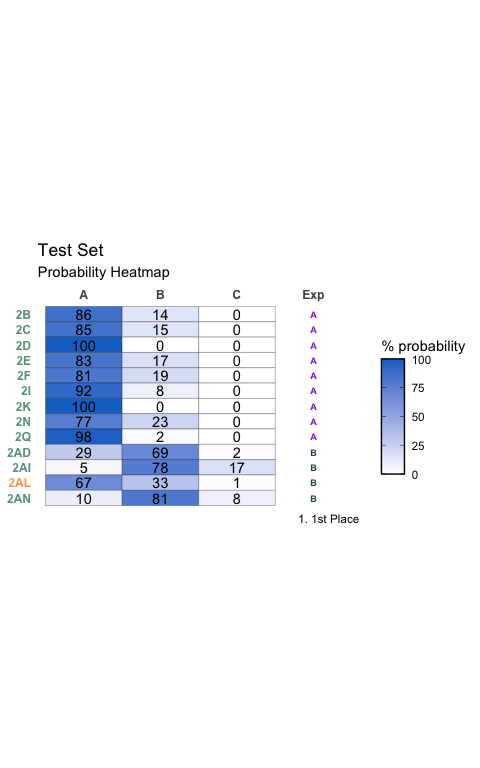
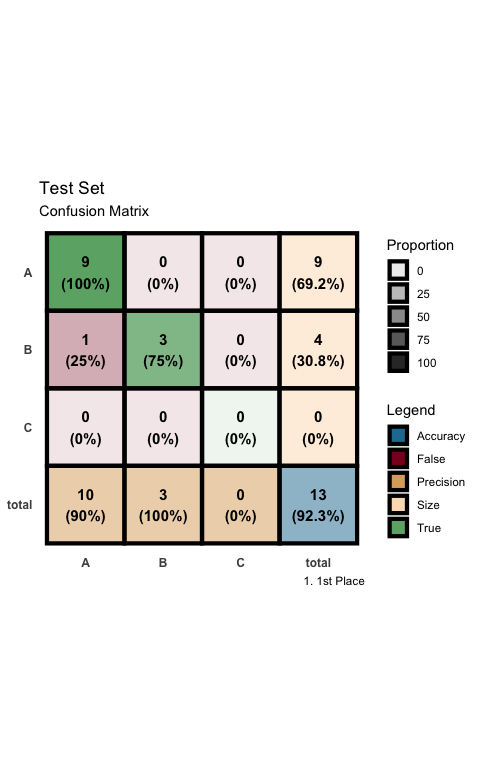
Full Model Stats - Overall Accuracy and Pseudo-R2

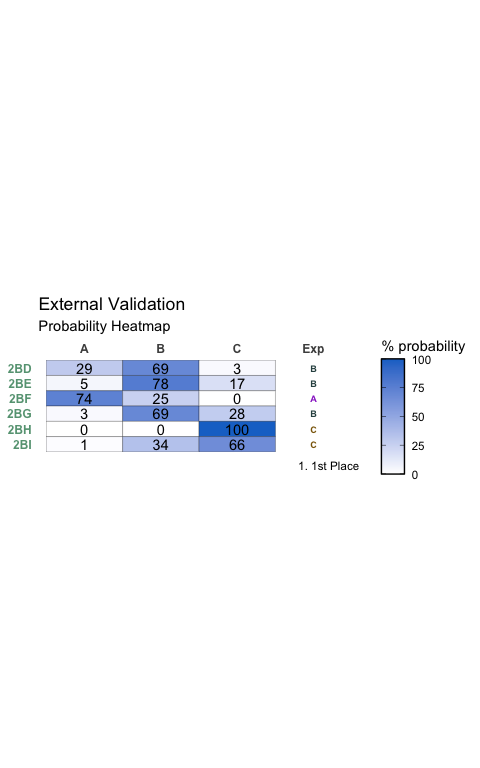
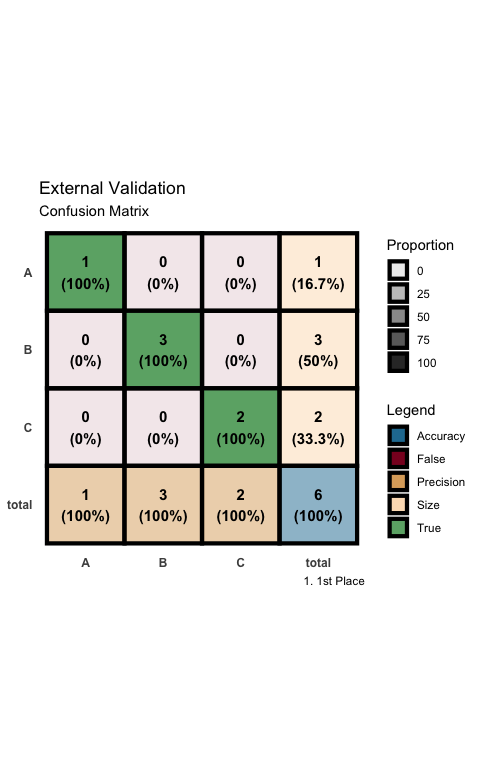
| Accuracy | McFadden\_R2 |
| --- | --- |
| 90.48% | 0.577 |

Model Coefficients

|  | x |
| --- | --- |
| -2-3- | -0.3801322 |
| dip\_y | -1.5680528 |
| Dist(2, 7) | 295.8494658 |
| NPA\_1 | 47.5647077 |







#### New Molecules Prediction

New Molecules - Prediction of Conditions

|  | Predicted | 1 | 2 | 3 |
| --- | --- | --- | --- | --- |
| 2BJ | 3 | 0 | 0 | 100 |
| 2BK | 1 | 73 | 27 | 0 |
| 2BL | 2 | 29 | 69 | 3 |
| 2BM | 1 | 100 | 0 | 0 |
| 2BN | 3 | 0 | 0 | 100 |
| 2BO | 3 | 0 | 0 | 100 |
| 2BP | 1 | 100 | 0 | 0 |
| 2BQ | 1 | 100 | 0 | 0 |
| 2BR | 1 | 100 | 0 | 0 |
| 2BS | 1 | 100 | 0 | 0 |
| 2BT | 1 | 100 | 0 | 0 |
| 2BU | 1 | 57 | 42 | 1 |
| 2BV | 2 | 25 | 72 | 3 |
| 2BW | 2 | 1 | 58 | 41 |

## Model Analysis - Variable Impact

### 1. NPA\_1 Removed

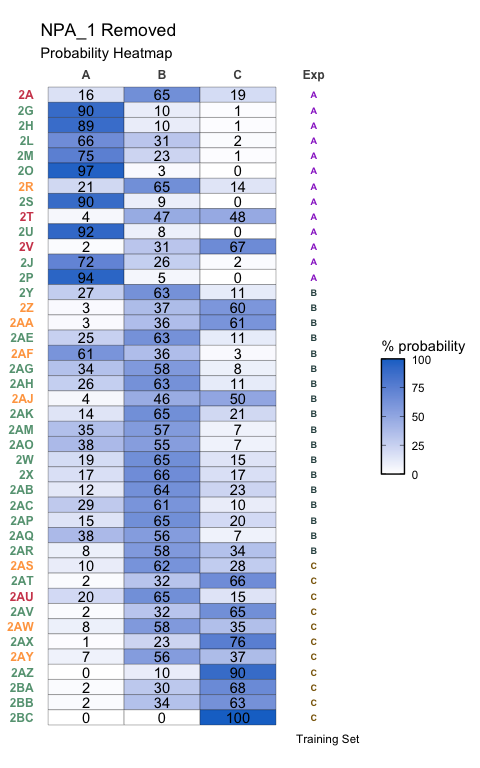
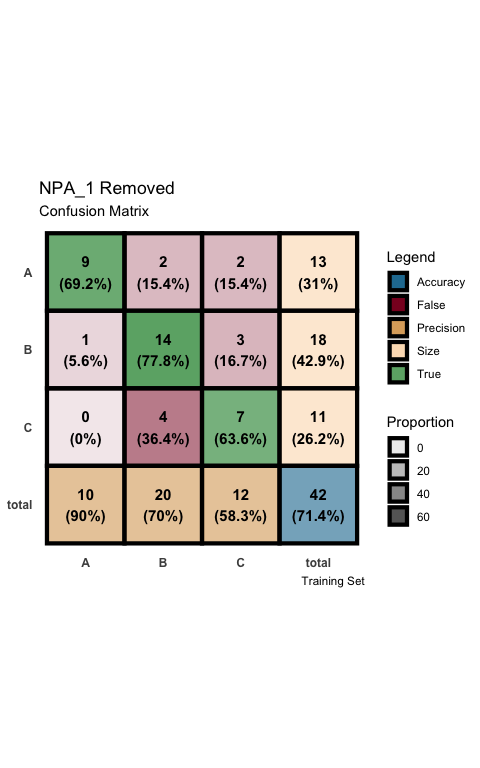
| formula |
| --- |
| class ~ -2-3- + dip\_y + Dist(2, 7) |

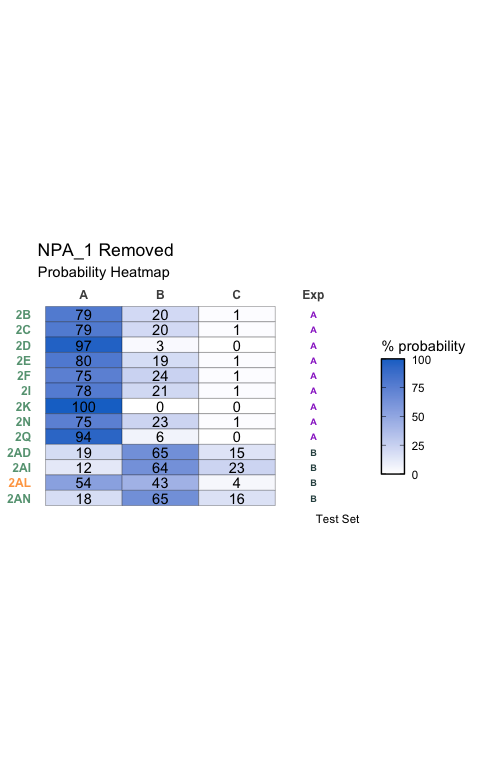
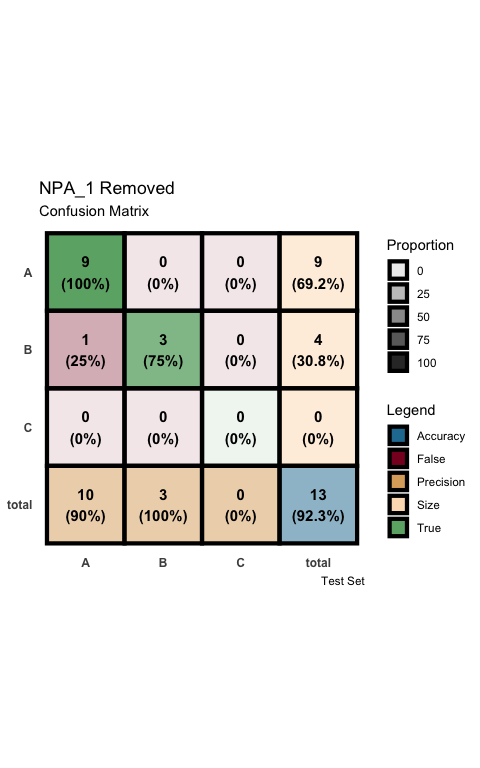
Full Model Stats - Overall Accuracy and Pseudo-R2

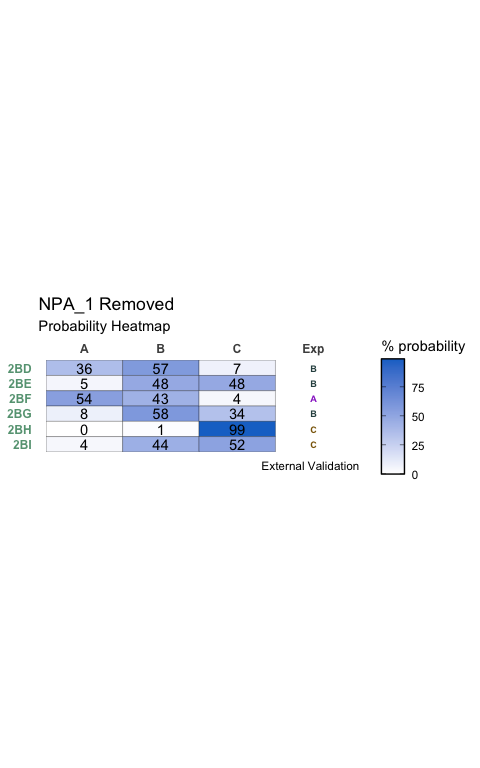
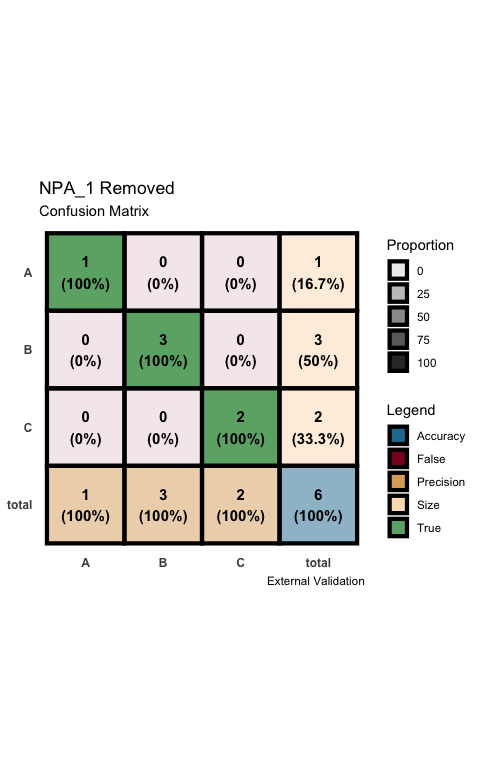
| Accuracy | McFadden\_R2 |
| --- | --- |
| 71.43% | 0.338 |

Model Coefficients

|  | x |
| --- | --- |
| -2-3- | -0.1882832 |
| dip\_y | -0.9381293 |
| Dist(2, 7) | 145.2808583 |







### 2. Dist(2, 7) Removed

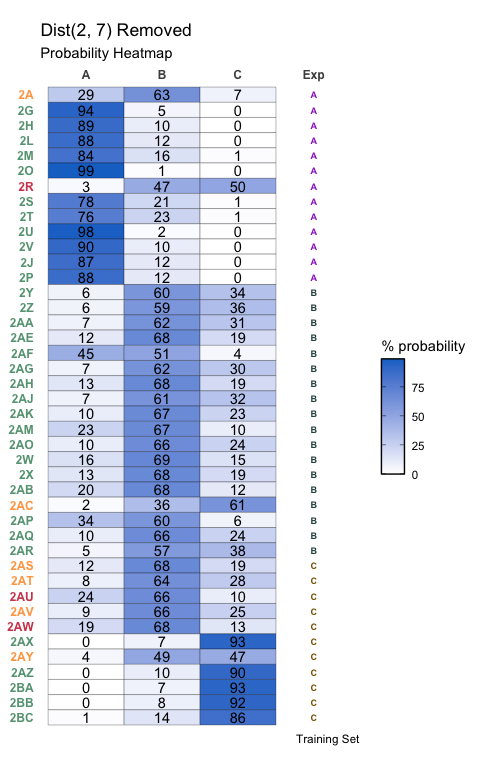
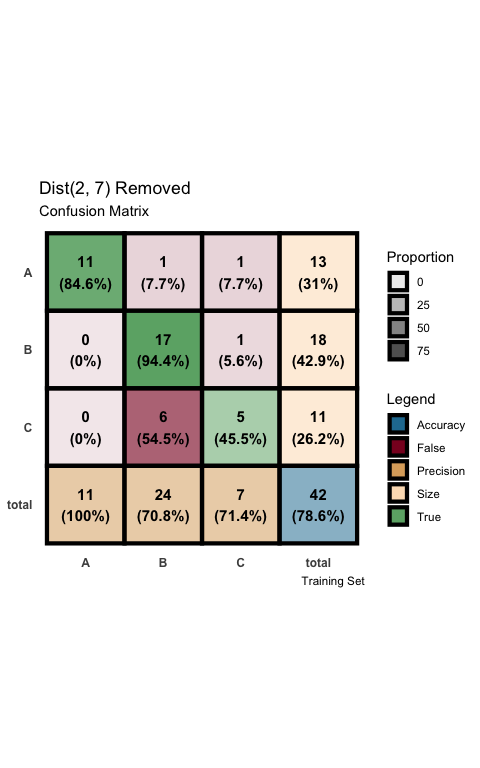
| formula |
| --- |
| class ~ -2-3- + dip\_y + NPA\_1 |

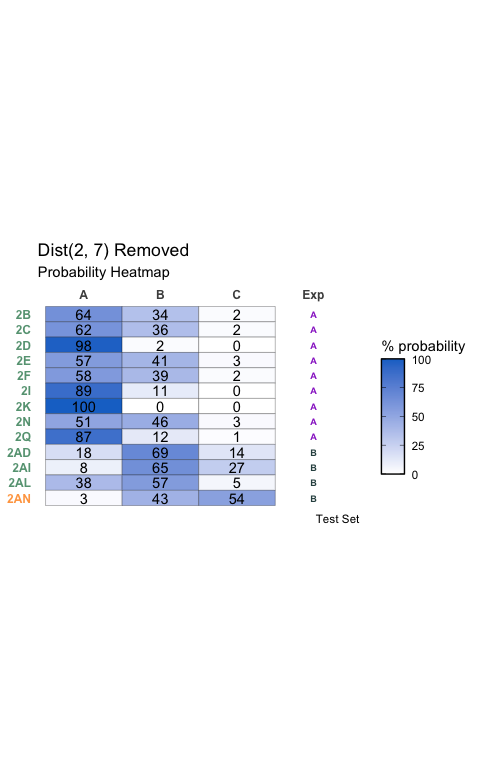
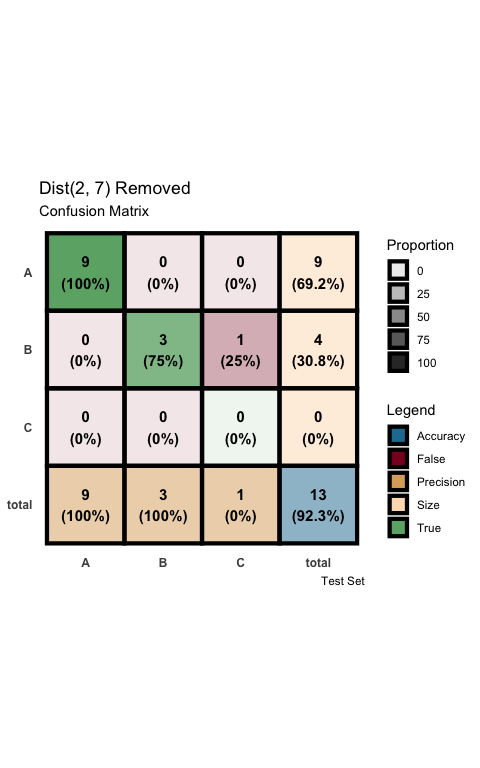
Full Model Stats - Overall Accuracy and Pseudo-R2

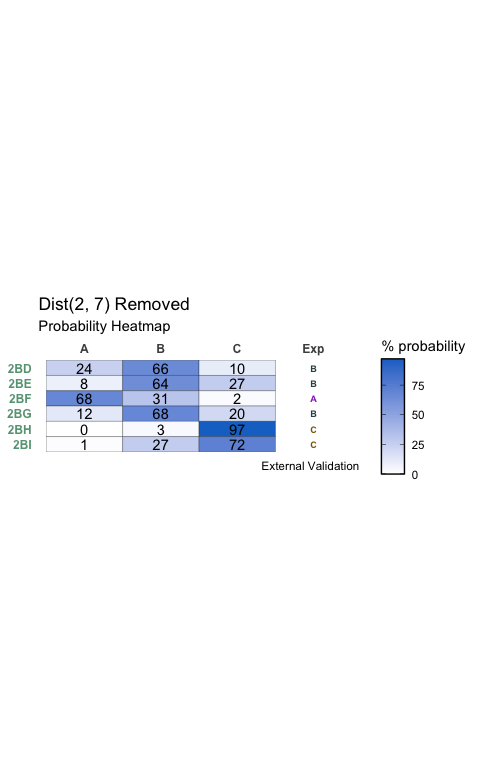
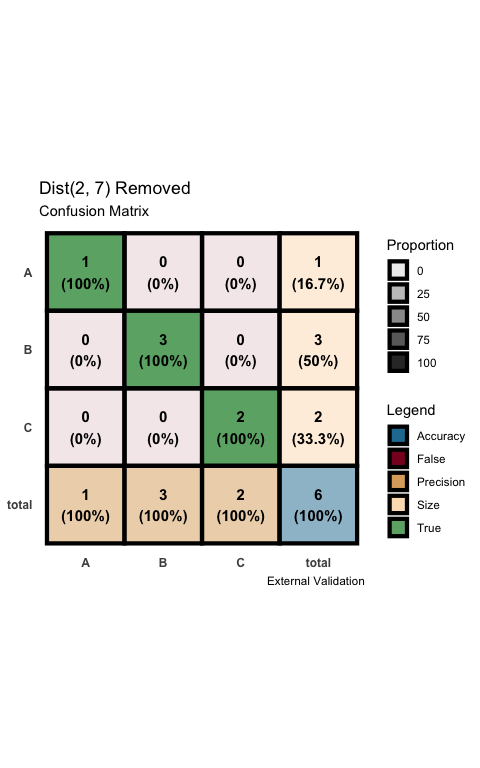
| Accuracy | McFadden\_R2 |
| --- | --- |
| 78.57% | 0.452 |

Model Coefficients

|  | x |
| --- | --- |
| -2-3- | -0.076705 |
| dip\_y | -1.830226 |
| NPA\_1 | 27.040639 |







### 3. dip\_y Removed

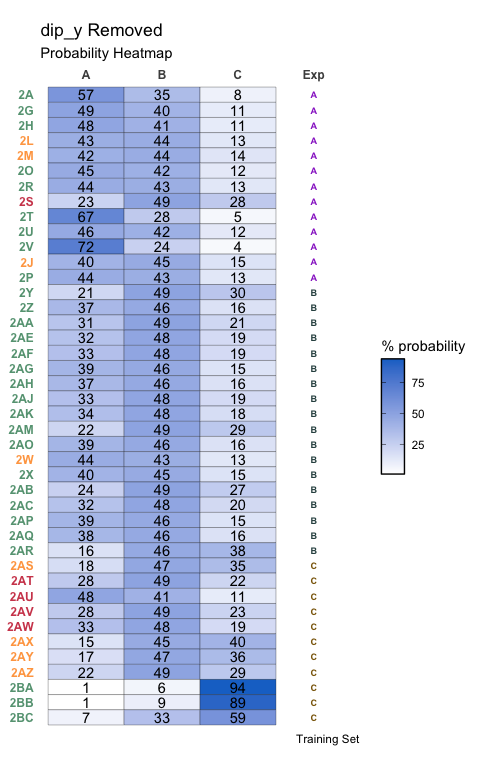
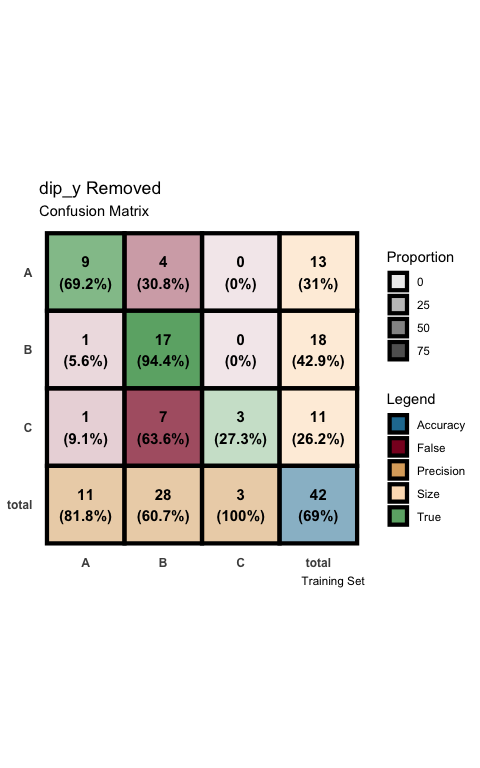
| formula |
| --- |
| class ~ -2-3- + Dist(2, 7) + NPA\_1 |

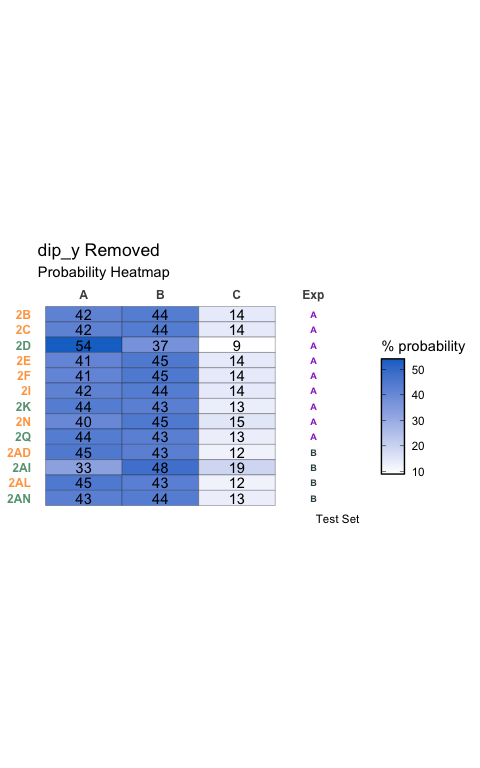
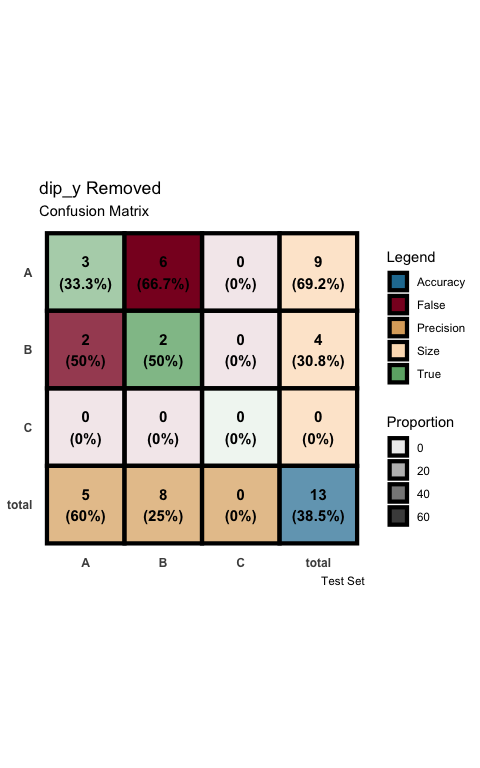
Full Model Stats - Overall Accuracy and Pseudo-R2

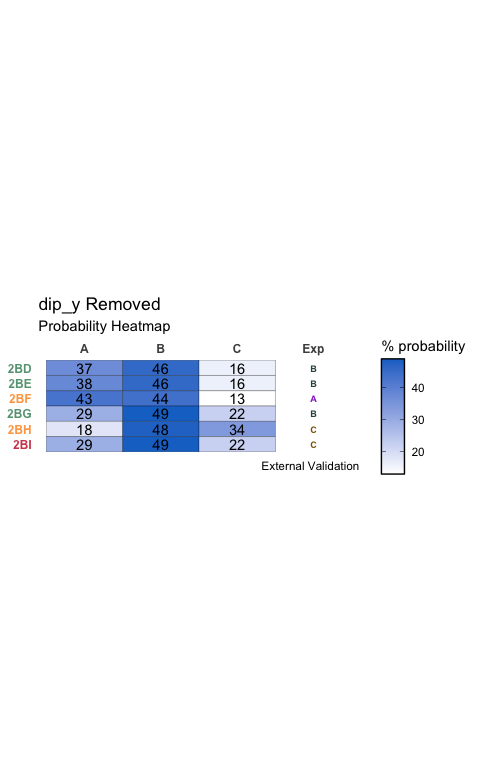
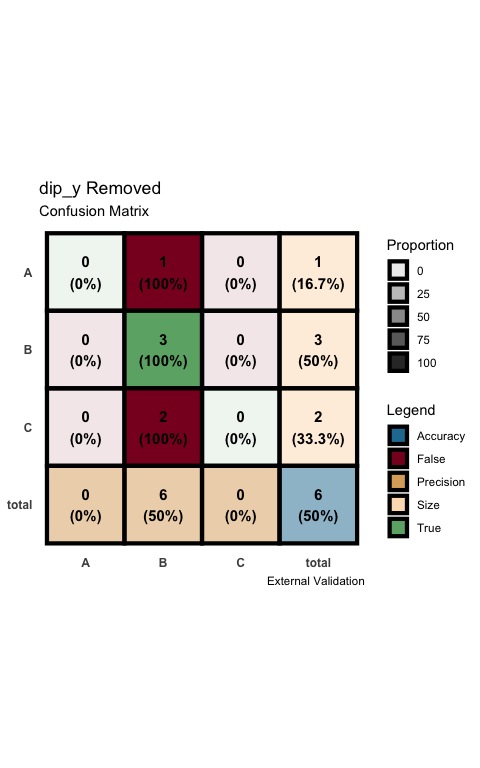
| Accuracy | McFadden\_R2 |
| --- | --- |
| 69.05% | 0.217 |

Model Coefficients

|  | x |
| --- | --- |
| -2-3- | -0.0879166 |
| Dist(2, 7) | 73.6506465 |
| NPA\_1 | 15.3035790 |







### 4. -2-3- Removed

| formula |
| --- |
| class ~ dip\_y + Dist(2, 7) + NPA\_1 |

Full Model Stats - Overall Accuracy and Pseudo-R2

| Accuracy | McFadden\_R2 |
| --- | --- |
| 76.19% | 0.344 |

Model Coefficients

|  | x |
| --- | --- |
| dip\_y | -1.932190 |
| Dist(2, 7) | -7.774941 |
| NPA\_1 | 23.524905 |

