**Logistic Regression Work Summary**

Methods Attempted:

1. Oversampling
   1. SMOTE
   2. SMOTEENN
   3. SMOTETomek
2. Regularization
   1. Ridge
   2. Lasso
   3. Elastic-Net
3. Scaling
   1. Standard Scaling
   2. MinMax Scaling
4. Feature Selection
   1. Lasso Regression
   2. Variance Threshold

Results Summary:

1. **GridSearch Determined Regularization Term**

|  |  |  |
| --- | --- | --- |
| Dataset | Coefficient Results | Model Performance |
| SMOTE |  | Accuracy = 0.73 |
| SMOTEENN |  | Accuracy = 0.69 |
| SMOTETomek |  | Accuracy = 0.71 |

1. **Attempting Lasso Regression for Feature Selection**

|  |  |  |
| --- | --- | --- |
| Dataset | Coefficient Results | Model Performance |
| SMOTE |  | Accuracy = 0.69 |
| SMOTEENN |  | Accuracy = 0.69 |
| SMOTETomek |  | Accuracy = 0.69 |

1. **Using Variance Threshold for Feature Selection**

|  |  |  |
| --- | --- | --- |
| Dataset | Coefficient Results | Model Performance |
| SMOTE |  | Accuracy = 0.67 |
| SMOTEENN |  | Accuracy = 0.65 |
| SMOTETomek |  | Accuracy = 0.65 |

1. **Attempting MinMax Scaling to make the model more sensitive to outliers**

|  |  |  |
| --- | --- | --- |
| Dataset | Coefficient Results | Model Performance |
| SMOTE |  | Accuracy = 0.92 |
| SMOTEENN |  | Accuracy = 0.92 |
| SMOTETomek |  | Accuracy = 0.92 |

1. **Frequency Summary**
   1. Frequencies Ranging from 4000 up to 4896 Hz -> Especially 4000, 4117, 4237
   2. Frequencies Ranging from 5495 up to 5993 HZ -> Especially 5495 and 5993
   3. Frequencies Ranging from 1834 up to 1887 Hz -> Especially 1834