|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **2 Classes with 15 Attributes and Census Data**  **Threshold 3** | | | | | |
| **Machine Learning Model** | **Accuracy Testing** | **Accuracy Training** | **Distribution** | **F-1 “Lower Performance”** | **F-1 “Higher Performance”** |
| MultiClass Classification Logistic Regression | 76.41% | 76.42% | * Lower Performance: 5834 * Higher Performance: 18905 | 0% | 87% |
| MultiClass Classification Logistic Regression –  **Naive Random Oversampling** | 60.24% (Balanced) | 61.94% (Balanced) | * Lower Performance: 14179 * Higher Performance: 14179 | 43% | 60% |
| MultiClass Classification Logistic Regression –  **SMOTE Oversampling** | 60.29% (Balanced) | 62.03% (Balanced) | * Lower Performance: 14179 * Higher Performance: 14179 | 43% | 61% |
| MultiClass Classification Logistic Regression –  **Cluster Centroids Undersampling** | 59.80% (Balanced) | 47.60% (Balanced) | * Lower Performance: 4375 * Higher Performance: 4375 | 43% | 51% |
| MultiClass Classification Logistic Regression –  **SMOTEEN, Combination of Over- and Under-Sampling** | 60.29% (Balanced) | 61.66% (Balanced) | * Lower Performance: 7176   Higher Performance: 5616 | 43% | 56% |
| Balanced Random Forest | 66.78% (Balanced) | 91.01% (Balanced) | * Lower Performance: 5834 * Higher Performance: 18905 | 49% | 74% |
| Easy Ensemble AdaBoost Classifier | 66.36% (Balanced) | 68.67% (Balanced) | * Lower Performance: 5834 * Higher Performance: 18905 | 48% | 73% |