# Model Evaluation Results Table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Method/Algorithm | Precision (P) | Recall (R) | Accuracy (A) | F1 Score (F) |
| Random Forest | Random Undersampling | 0.89 | 0.60 | 0.61 | 0.72 |
| SMOTE | 0.86 | 0.85 | 0.76 | 0.86 |
| ADASYN | 0.84 | 0.93 | 0.79 | 0.88 |
| Logistic Regression | Random Undersampling | 0.90 | 0.66 | 0.65 | 0.76 |
| SMOTE | 0.87 | 0.78 | 0.72 | 0.83 |
| ADASYN | 0.87 | 0.74 | 0.69 | 0.80 |
| XGBoost | Random Undersampling | 0.89 | 0.60 | 0.80 | 0.72 |
| SMOTE | 0.83 | 0.96 | 0.80 | 0.89 |
| ADASYN | 0.83 | 0.96 | 0.80 | 0.89 |
| GridSearch | 0.84 | 0.98 | 0.82 | 0.90 |
| Gradient  Boosting | SMOTE | 0.84 | 0.93 | 0.79 | 0.88 |
| Random Undersampling | 0.90 | 0.61 | 0.62 | 0.73 |
| ADASYN | 0.84 | 0.91 | 0.79 | 0.88 |
| GridSearch | 0.84 | 0.98 | 0.82 | 0.90 |

Given the significant class imbalance, where defaulters are the minority, we emphasized evaluating models using Precision, Recall, and F1 Score, rather than accuracy, to avoid misleading conclusions. Focusing on Class 1 (Defaulters) was critical because failing to identify defaulters can result in major financial losses. Recall was prioritized to minimize false negatives, as it was crucial to detect as many defaulters as possible. We used SMOTE, ADASYN, and Random Undersampling to handle the imbalance, with SMOTE and ADASYN showing superior performance, especially in recall.

Through GridSearchCV, we optimized the Gradient Boosting Classifier (GBC) and identified the best configuration, which included a learning rate of 0.2, max\_depth of 3, min\_samples\_leaf of 3, min\_samples\_split of 2, and n\_estimators set to 200. This tuned set of parameters helped maximize the model’s performance, particularly in terms of recall for identifying defaulters.

The tuned model achieved a precision of 84% and a recall of 98%, the highest recall among all models tested, demonstrating its effectiveness at correctly identifying defaulters. This high recall is critical in minimizing financial risk, and the balance of precision and recall highlights the model’s ability to make reliable predictions. The Gradient Boosting Classifier emerged as a strong and dependable model choice for our project.

We ran a similar GridSearchCV for XGBoost with ADASYN for class balancing, yielding results comparable to those of the Gradient Boosting Classifier. The XGBoost model also achieved high recall and precision, reinforcing its potential for detecting defaulters with similar efficacy, making it another reliable option for the task.