Model Training and Testing Report

We used a Multi-Model approach, in order to get the best model. Considering Binary classification problem **Logistic Regression** was our first go to model as it is easy to implement and fast as well. Since we need to make a decision on whether to provide a loan or not , our next models were **Decision Tree** and **Random Forest**. Also our data was fairly imbalanced and had missing values as well and so we thought to use **Gradient Boosting**.

1. **Logistic Regression**

   Training Accuracy Score : 81.47 %
   Testing Accuracy Score : 78.86 %
   Cross-Validation : 80.95 %

2. **Decision Tree**

   Training Accuracy Score : 100 %
   Testing Accuracy Score : 68.29 %
   Cross-Validation : 70.84 %

3. **Random Forest**

   Training Accuracy Score : 100 %
   Testing Accuracy Score : 78.05 %
   Cross-Validation : 78.83 %

4. **Gradient Boosting**

   Training Accuracy Score : 91.24%
   Testing Accuracy Score : 90.24%

As we can see, the **Gradient Boosting** model has the highest accuracy and thus is the best model among the four.