**Loan Status Prediction**

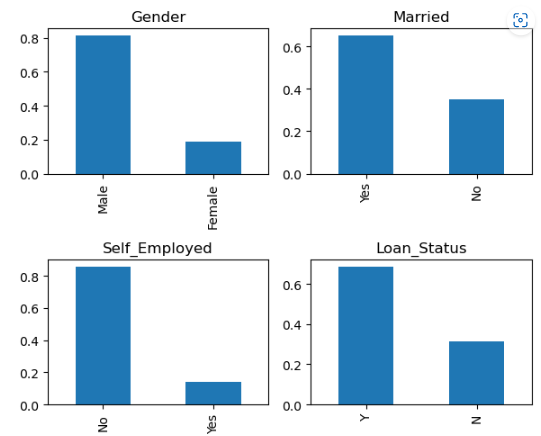
**Objective**: To build a Machine Learning Model to predict the loan to be approved or to be rejected for an applicant based on previous loan approval data.

**Model Implementation**: Lazy Predict, Logistic Regression, Random Forest Classifier, Grid Search CV

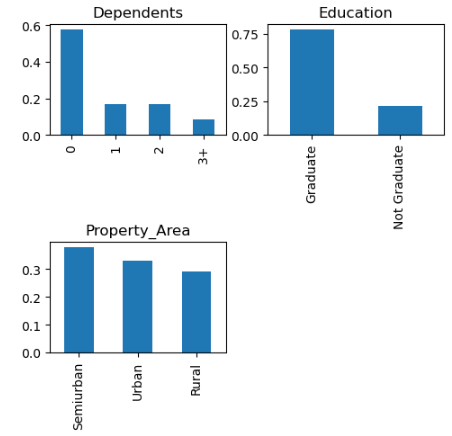
**Model Accuracy**: **83.78%**

**Key Insights:**

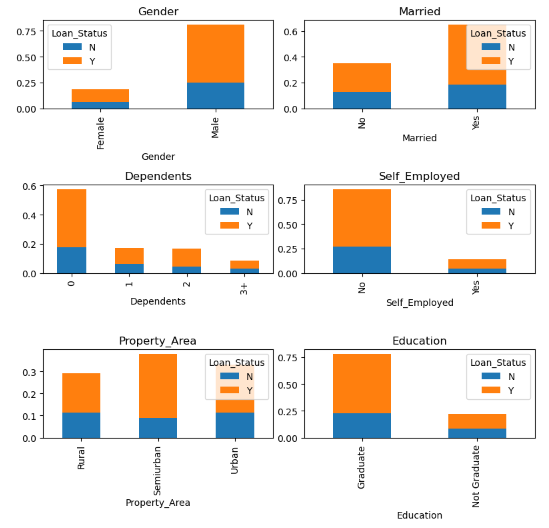
By visualizing the **Categorical Features** of the data, It shows that most of the individuals in the dataset are males, and the majority of them are married and not self-employed. Additionally, around 65% of the loan applications have been approved. These observations can help in understanding the dataset and guide further analysis.



At the same time, when we look into **Ordinal Features**, the data suggests that many applicants have no dependents, most are graduates, and live in semi-urban areas. These observations can provide insights into the dataset and guide further analysis.

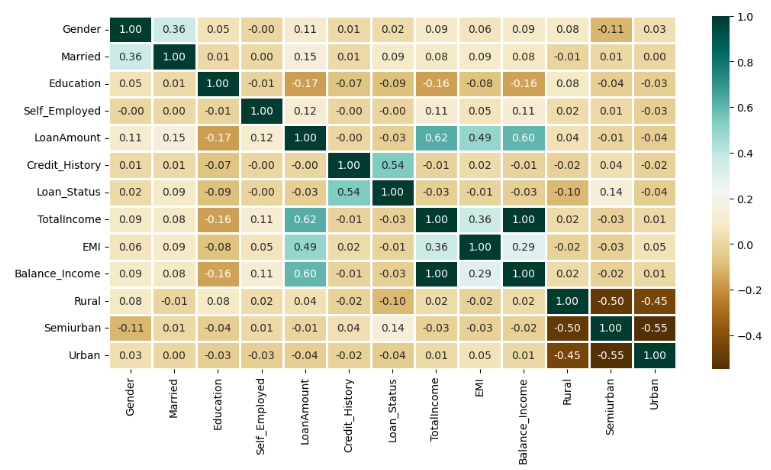


When we compare the **Categorical features with the target variable**,

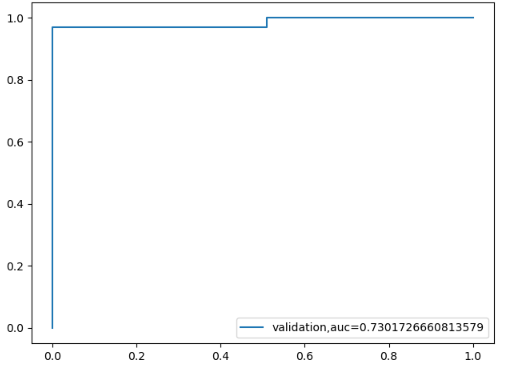


The data shows that loan approval rates are similar between genders. Applicants with fewer dependents, who are married, graduates, and from semi-urban areas have a higher chance of getting their loan approved.

From the below heat map of the correlation of the features, it is clear that credit history has a huge impact on the loan status. Furthermore, some of the independent features as well correlate with each other which is visualized in the below plot



**ROC Curve**



The confusion matrix shows that the classifier has an 84% accuracy and performs better in predicting class 1 than class 0, with higher precision, recall, and F1-scores for class 1. The low recall score for class 0 suggests that the classifier needs improvement in identifying examples belonging to class 0. Overall, there is room for improvement in the classifier's performance.

**Precision** **Recall**  **f1-score** **Support**

0 0.86 0.49 0.62 51

1 0.83 0.97 0.90 134

**Accuracy**  0.84 185

**macro avg** 0.85 0.73 0.76 185

**weighted avg**  0.84 0.84 0.82 185

**Conclusion**:

Based on these metrics, we can say that the classifier performs better in predicting class 1 than class 0, as the precision, recall, and F1-score for class 1 are higher than those for class 0, which is crucial as banks need to be careful on who they are going to give a loan than rejecting.Hence, In order to minimize the risk of loan defaults, it may be more cautious in giving loans to applicants in class 0 with lower precision and recall scores.