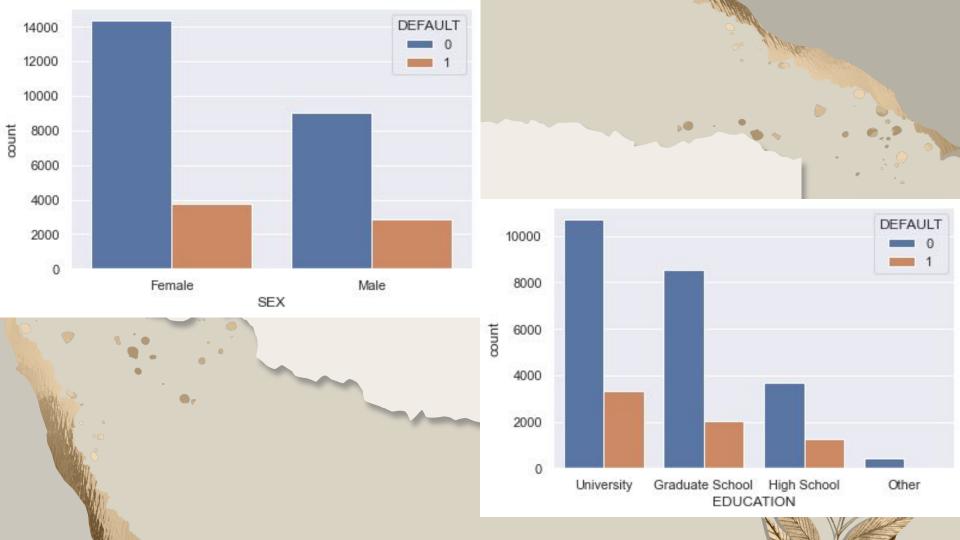


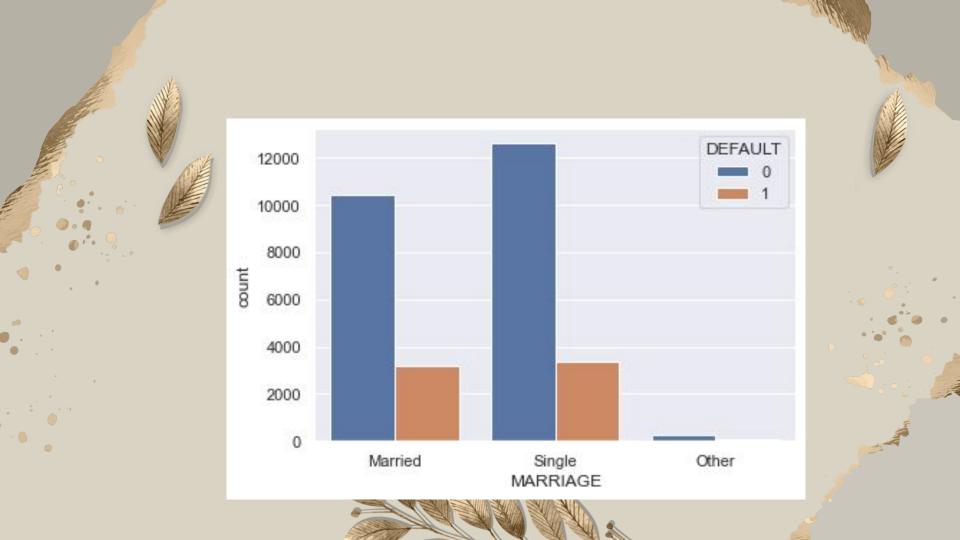
Why is it important?

- Limit Financial Growth of the company
- Decrease in Revenue
- Decrease in Profit
- Decrease in Liquidity
- Decrease in Lending Potential
- Loss of Confidence from Depositors & Investors

Features

- Amount of Credit
- Gender
- Education
- Marital Status
- Age
- History of Past Payments
- Amount of Past Bills





Models and Metrics

Model	F2 Score	Accuracy	Comments
K Nearest Neighbor	.3915	.8091	GridSearchCV to optimize hyper parameter
Logistic regression	.3696	.8163	Grid Search CV to optimise C
Logistic Regression (Threshold = .4)	.5903	.6681	Imbalance Class Treatment: Balanced Class Weight & Threshold Grid Search CV to optimise C

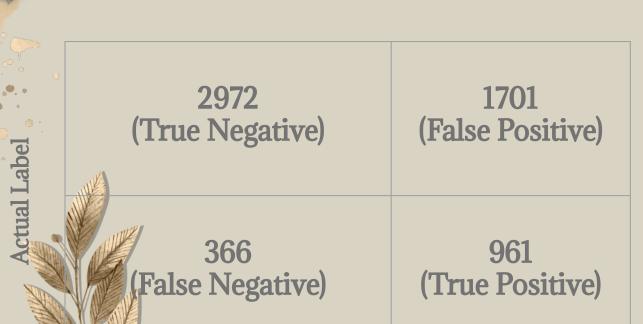




Models and Metrics

Model	F2 Score	Accuracy	Comments
Random Forest (Threshold = .17)	.588	.233	Imbalance Class Treatment: Threshold Grid Search CV to optimise depth and estimators
Random Forest (Threshold = .17)	.602	.655	Feature Engineering Imbalance Class Treatment: Threshold & Balanced class weight Grid Search CV to optimise depth and estimators

CONFUSION MATRIX



Predicted Label



RECALL = .72

F2 SCORE = .60



CONCLUSION:

- False Negative are more harmful than False Positive
- 72.4 % of customers who will default were correctly identified
- 36.5 %(FP) of customers who will not default were in-correctly identified
- 1% of incorrectly classified non defaulting customers (FP) may switch companies because of continuous follow up by companies.
- 72% (TP) were followed up and offered alternate plans of payment
- Gender, Education and timely payment of past dues are the most important factors to predict default probability in future

FEATURE IMPORTANCE



