

Personal Loan Acceptance Analysis

Synopsis: The purpose of this analysis is to determine the ideal customer demographic for Universal Bank. This analysis will consist of using 5,000 previous customer information stored in a database and using that information to determine the type of customer that is most likely to accept a personal loan. A logistic regression model will be used to identify the significant factors that will find the correct target market for a potential customer that will accept a personal loan and we will use summary statistics like odd ratios to determine who to target for the marketing campaign.

Regression Analysis: For this marketing analysis, the logistic regression is best suited. A logistic regression is a statistical machine learning technique that fits data into binary classification which predicts probabilities using a logit function. Here, several categories of customer information will be taken into consideration including but not limited to mortgage, college standings, securities, and credit cards. These variables will be transformed into binary outputs with 1 indicating they possess the characteristic and a 0 if they do not. Once they are transformed, the information will be processed in python. In python, sklearn library will be used to train and predict the model using the logistic regression. From there, python will construct the summary results of our model with the three in focus being the confusion matrix, the classification report, and the odds ratio.

Results: The following tables below show the results of our confusion matrix, classification report, and the odds ratio:

Odds:		Confusion Matrix:					
	0	[[2200 44]					
Family	2.293521	[135 121]]					
CCAvg	1.461246	Accuracy: 0.9284					
EducProf	1.322243	Classification Report:					
EducGrad	1.303714		precision	recall	f1-score	support	
CD	1.278414						
Securities	1.047099						
Income	1.035658						
CreditCard	1.017986	0	0.94	0.98	0.96	2244	
Experience	1.013019	1	0.73	0.47	0.57	256	
Online	1.005519						
Mortgage	1.001034						
CustomerID	0.999942	accuracy			0.93	2500	
ZIPCode	0.999901	macro avg	0.84	0.73	0.77	2500	
Age	0.999573	weighted avg	0.92	0.93	0.92	2500	
EducUnder	0.576542						

The confusion matrix shows numerically how the model performed. The left diagonal shows the numbers the model was correctly able to identify and the right diagonal shows the incorrect numbers identified. As per the matrix, it shows we had 2200 true negatives, 44 false positives, 135 false negatives, and 121 true positives. With an accuracy percentage of 0.9284, the model performed well.

The classification report includes the following statistics: precision, recall, f1-score, and support. The Precision is the ratio of correctly predicted positive observations to the total predicted positive observations, recall is the ratio of correctly predicted positive observations to all observations in actual positive class, and f1 score is the mean of the previous two categories. In all of the instances, people who were not part of a particular category were more accurate than the ones who were.

Finally, the odds ratio represents how a unit change in a predictor variable affects the odds of the outcome. If the odds are above 1, that means each increase in the predictor variable increases the odds of that event happening. If it's close to 1, that means each increase in the predictor variable does not really change the odds of that even happening. If it's less than 1, that means each increase in the predictor variable decreases the odds of that occurrence. To capture whether the predictor variable is significant or not we will use a significance level of ± 0.05 from 1. After doing that, the significant variables in our analysis are: Family, CCAvg, EducProf, EducGrad, CD, and EducUnder.

Conclusion: To be able to best target individuals who are likely to accept a loan, Universal Bank should target individuals who:

- Have a family
- Use Credit Cards
- Have a Professor or Masters degree
- Have a CD in their bank accounts

However, they should avoid people with just an undergraduate degree as they are not likely to be interested in the loans.