greatlearning
Power Ahead

AllLife Bank Project

Contents

- Business Problem Overview
- Data Overview
- Exploratory Data Analysis (EDA)
- Model Performance Summary
- Business Insights and Recommendations

Business Problem Overview and Solution Approach

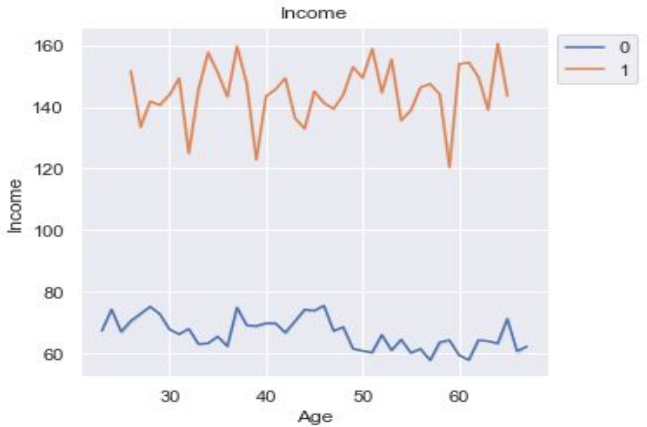
- AllLife is a US bank has a growing customer base and wants to capitalize on it by converting its liability customers into personal loan customers.
- Bank would profit from converting such customers to assets by earning interest on loans.
- Previous campaigns showed a healthy conversion rate of over 9% success which bank wants to improve further by targeting the right customers hence increasing the success ratio.
- We will develop classification models (Logistic Regression model and Decision Tree model) by analyzing data related to the various attributes of a customers (Age, Income, Education, etc.) to help the marketing team identify the potential customers who would be interested in taking a personal loan and use it to devise profitable strategies.

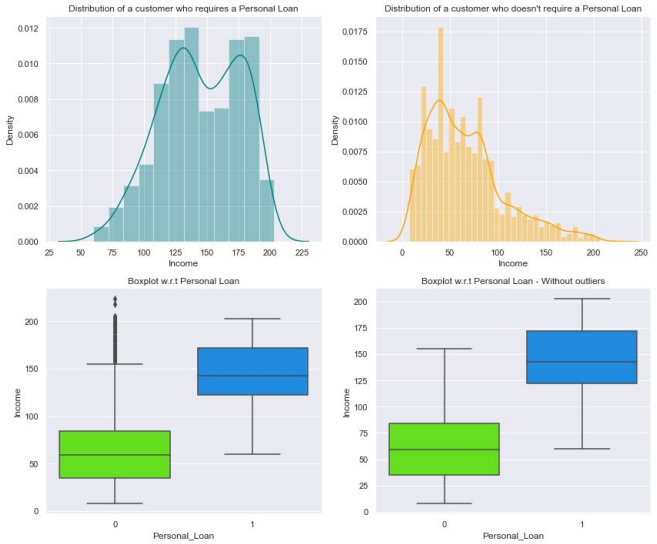
Data Overview

- The data contains information about 5000 customers.
- The information include Age, Income, years of experience they have, power of the car, Annual income of the customer, Avg. spending on credit cards per month , Family size of the customer, Education Level, Value of house mortgage, and more.
- Experience column has some negative values, which warrants for a anomaly check.
- We will extract first two digits from the ZIPCode column as first digit indicates the regions and second digit indicates the sub regions(States) to reduce the number of unique values for our analysis.

Exploratory Data Analysis

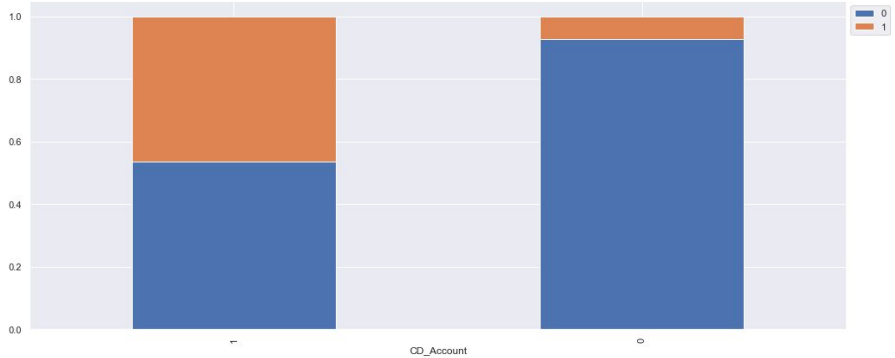
- Financial stability plays a big part in the requirement of Loan.
- Those customers who have an income higher than 100k dollars are the potential customers who will take the Personal Loan.

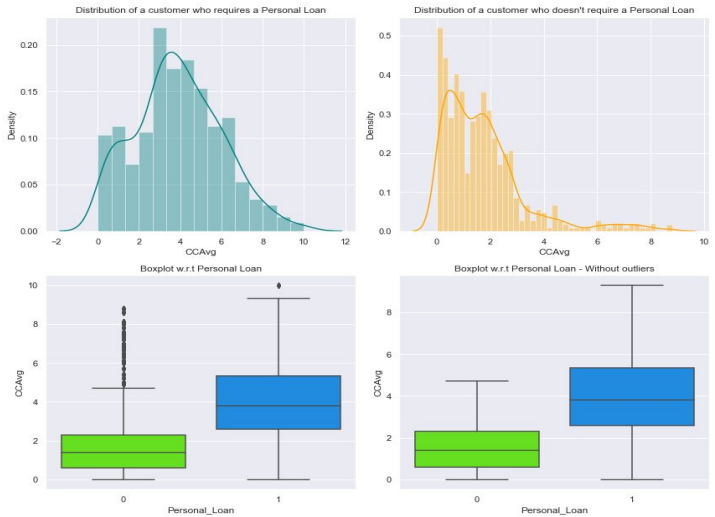




Exploratory Data Analysis

- ~50% of the customers who have a certificate of deposit with the bank (CD_Account) are the ones that have the requirement of a Personal Loan.

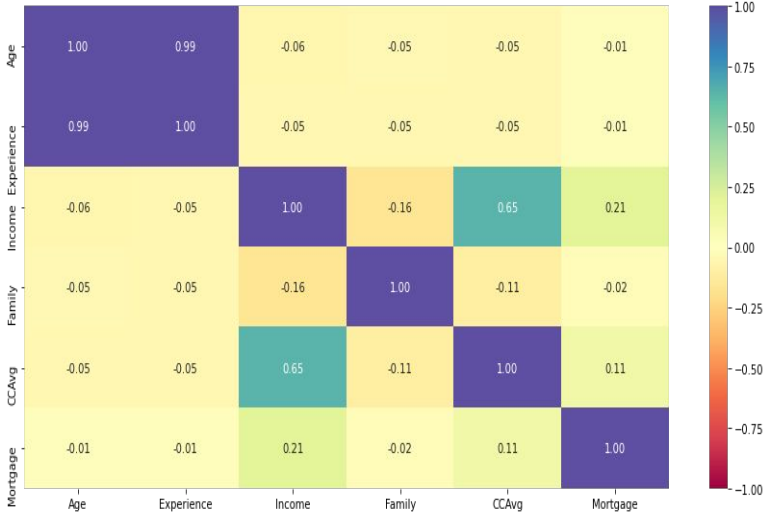




- The customers with a CCAvg of greater than 5 can be the potential customers who opt for Personal Loan.

Exploratory Data Analysis

- Age and Experience have a perfect correlation hence one of these variables can be dropped while model building as they will provide the same information to the model.
- Income and CCAvg has a moderate correlation which makes sense as the Income increases the spendings might also increase.
- Family has a negative correlation with Income, which is quite surprising as family size increases the income of the family decreases. But this correlation is not too strong to conclude.

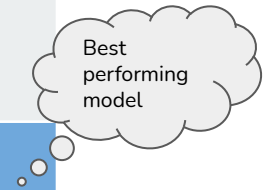


Model Performance Summary

- We want to predict whether a customer will take a personal loan or not using their information initially provided to us.
- We will use Recall as the performance metric for our model.
- Predicting a customer will take the personal loan and the customer doesn't take the loan will result in loss of opportunity and bank would want to minimize such false negatives.
- The insights we get from our Logistic Regression are:
  - Income, Family, Education, CCAvg and CD_Account are the important predictors.
  - Online, CreditCard, Securities_Account, ZIPCode have very little importance.
  - 1 unit change in the Family will increase the odds of taking a personal loan by 91.25%
  - 1 unit change in CCAvg will increase the odds of taking a personal loan by 22.16%
  - 1 unit change in Income will increase the odds of taking a personal loan by 6.11%
  - Customers with Income less than 116.5k dollars, CCAvg less than 2.95, and Income less than 106.5k dollars have fewer chances of taking a Personal Loan.
  - Customers with a family size greater than 2, and education level more than undergraduate has more chances of buying a loan..

Model Performance Summary

Model	Train Accuracy	Test Accuracy	Train Recall	Test Recall	Train Precision	Test Precision
Logistic regression	0.95	0.95	0.67	0.65	0.85	0.89
Logistic Regression (Optimal threshold)	0.91	0.90	0.88	0.88	0.56	0.47
Decision Tree	1	0.98	1	0.91	1	0.95
Decision Tree - Pre Pruning	0.98	0.97	0.87	0.78	0.99	1
Decision Tree - Post Pruning	1	0.98	1	0.91	1	0.95



Business Insights and Recommendations

- We should target higher income customers - Customers having income above 116k dollars should the leads.
- We should offer personal loans to customers who are highly education level greater than undergraduate as they have a higher chance of taking a personal loan.
- We should target the customers with a big family size as they would me more interested in taking a personal loan -Generally a family size of or more than 3 members have more chances of a customer taking a personal loan.
- We should target customers who have the certificate of deposit with the bank (CD_Account) - Our analysis shows that ~50% of the customers have a requirement of Personal Loan.

greatlearning
Power Ahead

Happy Learning !