

Mini Project | Supervised Learning and Ensembles

[Submit Assignment](#)

Due 28 Jan by 23:00 **Points** 50 **Submitting** a file upload **Available** 14 Jan at 8:00 - 28 Jan at 23:00 15 days

Dear Participant,

The Supervised Learning and Ensembles Lab is for 50 marks. It is due on 28th January 2018, 11:00 pm. Kindly submit it before the deadline.

Case Study-

Campaign for selling personal loans

This case is about a bank (Thera Bank) which has a growing customer base. Majority of these customers are liability customers (depositors) with varying size of deposits. The number of customers who are also borrowers (asset customers) is quite small, and the bank is interested in expanding this base rapidly to bring in more loan business and in the process, earn more through the interest on loans. In particular, the management wants to explore ways of converting its liability customers to personal loan customers (while retaining them as depositors). A campaign that the bank ran last year for liability customers showed a healthy conversion rate of over 9% success. This has encouraged the retail marketing department to devise campaigns with better target marketing to increase the success ratio with minimal budget.

The department wants to build a model that will help them identify the potential customers who have higher probability of purchasing the loan. This will increase the success ratio while at the same time reduce the cost of the campaign.

The file given below contains data on 5000 customers. The data include customer demographic information (age, income, etc.), the customer's relationship with the bank (mortgage, securities account, etc.), and the customer response to the last personal loan campaign (Personal Loan). Among these 5000 customers, only 480 (= 9.6%) accepted the personal loan that was offered to them in the earlier campaign.

Data Set - [Bank Personal Loan Modelling-1.xlsx](#) 

Considering the information provided above, follow the steps given below --

1. Read the column description and ensure you understand each attribute well.
2. Study the data distribution in each attribute, share your findings.
3. Get the target column distribution. Your comments.
4. Split the data into training and test set in the ratio of 70:30 respectively.
5. Use a classification model to predict the likelihood of a liability customer buying personal loans.
6. Explain why you chose one model over the other (do not use ensemble techniques yet).
7. Use ensemble techniques to improve the performance.

Please Note:

- Total marks allotted for Mini Project is 50.
- **Only one person per group** should make the submission for their group.
- Please mention the group number and members on the first page of the submission.
- Please submit working code and output (of each step) along with it in pdf,html and ipynb format.
- Please add necessary comments in all the files and make a managerial report based on that.
- Name format of files should always be **Group #_PGPBDA.B.AUG16_SLE_MiniPro.extension**

Marks Distribution is as follows --

1. Step 1 : Nil Points
2. Step 2 : 10 points
3. Step 3 : 5 points
4. Step 4 : 5 points
5. Step 5 : 10 points
6. Step 6 : 10 points
7. Step 7 : 10 points