G **Great Learning**

← Go Back to Advanced Machine Learning

≔ Course Content

Credit Card Users Churn Prediction : Problem Statement

Submission type	:	File Upload
Due Date	:	Jun 22, 4:30 AM MST
Total Marks	:	60
Available from	:	May 30, 7:30 AM

Description ⌃

Background & Context

The Thera bank recently saw a steep decline in the number of users of their credit card, credit cards are a good source of income for banks because of different kinds of fees charged by the banks like annual fees, balance transfer fees, and cash advance fees, late payment fees, foreign transaction fees, and others. Some fees are charged to every user irrespective of usage, while others are charged under specified circumstances.

Customers' leaving credit card services would lead the bank to loss, so the bank wants to analyze the data of customers and identify the customers who will leave their credit card services and the reason for same – so that the bank could improve upon those areas

You as a Data Scientist at Thera Bank need to come up with a classification model that will help the bank improve its services so that customers do not renounce their credit cards

You need to identify the best possible model that will give the required performance

Objective

1. Explore and visualize the dataset.

2. Build a classification model to predict if the customer is going to churn or not

3. Optimize the model using appropriate techniques

4. Generate a set of insights and recommendations that will help the bank

Data Dictionary:

- CLIENTNUM: Client number. Unique identifier for the customer holding the account

- Attrition_Flag: Internal event (customer activity) variable - if the account is closed then "Attrited Customer" else "Existing Customer"

- Customer_Age: Age in Years

- Gender: Gender of the account holder

- Dependent_count: Number of dependents

- Education_Level: Educational Qualification of the account holder - Graduate, High School, Unknown, Uneducated, College(refers to a college student), Post-Graduate, Doctorate.

- Marital_Status: Marital Status of the account holder

- Income_Category: Annual Income Category of the account holder

- Card_Category: Type of Card

- Months_on_book: Period of relationship with the bank

- Total_Relationship_Count: Total no. of products held by the customer

- Months_Inactive_12_mon: No. of months inactive in the last 12 months

- Contacts_Count_12_mon: No. of Contacts between the customer and bank in the last 12 months

- Credit_Limit: Credit Limit on the Credit Card

- Total_Revolving_Bal: The balance that carries over from one month to the next is the revolving balance

- Avg_Open_To_Buy: Open to Buy refers to the amount left on the credit card to use (Average of last 12 months)
- Total_Trans_Amt: Total Transaction Amount (Last 12 months)
- Total_Trans_Ct: Total Transaction Count (Last 12 months)
- Total_Ct_Chng_Q4_Q1: Ratio of the total transaction count in 4th quarter and the total transaction count in 1st quarter
- Total_Amt_Chng_Q4_Q1: Ratio of the total transaction amount in 4th quarter and the total transaction amount in 1st quarter
- Avg_Utilization_Ratio: Represents how much of the available credit the customer spent

Submission Guidelines

1. There are two ways to work on this project:

i. Full-code way: The full code way is to write the solution code from scratch and only submit a final Jupyter notebook with all the insights and observations.

ii. Low-code way. The low-code way is to use an existing solution notebook template to build the solution and then submit a business presentation with insights and recommendations.

The primary purpose of providing these two options is to allow learners to opt for the approach that aligns with their individual learning aspirations and outcomes. The below table elaborates on these two options.

Submission type	Who should choose	What is the same across the two	What is different across the two	Final submission file [IMP]	Submission Format
Full-code	Learners who aspire to be in hands-on coding roles in the future focussed on building solution codes from scratch	Perform exploratory data analysis to identify insights and recommendations for the problem	Focus on code writing: 10-20% grading on the quality of the final code submitted	Solution notebook from the full-code template submitted in .html format	.html
Low-code	Learners who aspire to be in managerial roles in the future-focussed on solution review, interpretation, recommendations, and communicating with business		Focus on business presentation: 10-20% grading on the quality of the final business presentation submitted	Business presentation in .pdf format with problem definition, insights, and recommendations	.pdf

Please follow the below steps to complete the assessment. Kindly note that if you submit a presentation, ONLY the presentation will be evaluated. Please make sure that all the sections mentioned in the rubric have been covered in your submission.

i. Full-code version

- Download the full-code version of the learner notebook.

- Follow the instructions provided in the notebook to complete the project.

- Clearly write down insights and recommendations for the business problems in the comments.

- Submit only the solution notebook prepared from the learner notebook [format: .html]

ii. Low-code version

- Download the low-code version of the learner notebook.

- Follow the instructions provided in the notebook to complete the project.

- Prepare a business presentation with insights and recommendations to the business problem.

- Submit only the presentation [format: .pdf]

2. Any assignment found copied/plagiarized with other submissions will not be graded and awarded zero marks.

3. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.

4. Submission will not be evaluated if

- it is submitted post-deadline, or,

- more than 1 file is submitted.

Best Practices for Full-code submissions

- The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.

- The notebook should be run from start to finish in a sequential manner before submission.

- It is important to remove all warnings and errors before submission.
- The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb).
- Please refer to the FAQ page for common project-related queries.
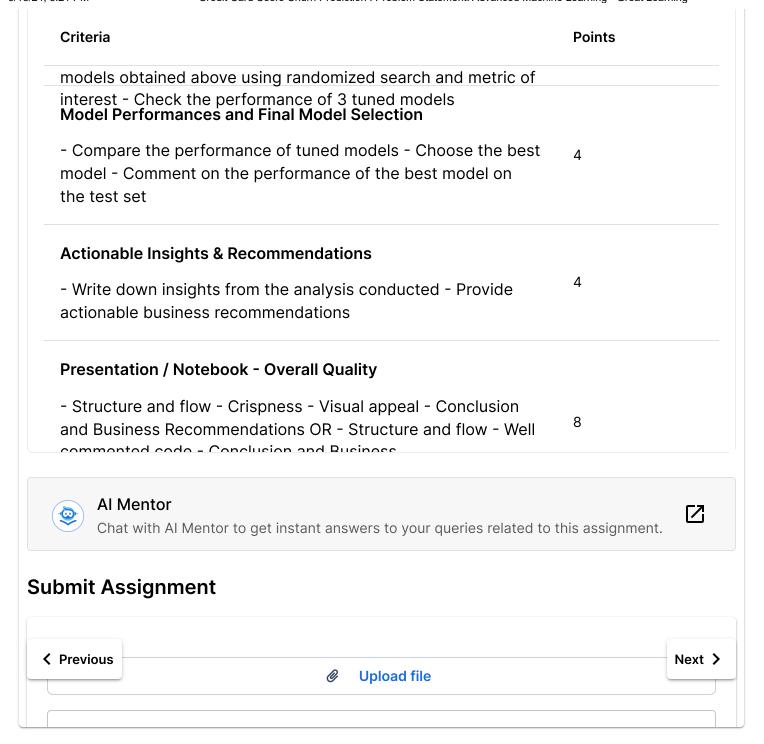
Best Practices for Low-code submissions

- The presentation should be made keeping in mind that the audience will be the Data Science lead of a company.
- The key points in the presentation should be the following:
  - Business Overview of the problem and solution approach
  - Key findings and insights which can drive business decisions
  - Business recommendations
  - Focus on explaining the key takeaways in an easy-to-understand manner.
  - The inclusion of the potential benefits of implementing the solution will give you the edge.
- Copying and pasting from the notebook is not a good idea, and it is better to avoid showing codes unless they are the focal point of your presentation.
- The presentation should be submitted as a PDF file (.pdf) and NOT as a .pptx file.
- Please refer to the FAQ page for common project-related queries.

Happy Learning!

Scoring guide (Rubric) - Credit card Users Churn Prediction   ⌃

Criteria	Points
Exploratory Data Analysis and Insights	8

Criteria	Points
- Problem definition, questions to be answered - Data background and contents - Univariate analysis - Bivariate analysis - Key meaningful observations on individual variables and the relationship between variables	
Data pre-processing - Prepare the data for analysis - Feature Engineering - Missing value Treatment - Outlier Treatment Note: Please ensure no data leakage occurs among train-test and validation sets	5
Model Building - Original Data - Build 5 models (from decision trees, bagging and boosting methods) - Comment on the model performance * You can choose NOT to build XGBoost if you are facing issues with the installation	6
Model building - Oversampled data - Oversample the train data - Build 5 models (from decision trees, bagging and boosting methods) - Comment on the model performance * You can choose NOT to build XGBoost if you are facing issues with the installation	6
Model building - Undersampled data - Undersample the train data - Build 5 models (from decision trees, bagging and boosting methods) - Comment on the model performance * You can choose NOT to build XGBoost if you are facing issues with the installation	6
Model Performance Improvement using Hyperparameter Tuning - Choose models that might perform better after tuning (tune at least 3 models out of 15 built in the previous steps) - Provide proper reasoning for tuning that model - Tune the best 3	13

Criteria	Points
models obtained above using randomized search and metric of interest - Check the performance of 3 tuned models **Model Performances and Final Model Selection** - Compare the performance of tuned models - Choose the best model - Comment on the performance of the best model on the test set	4
Actionable Insights & Recommendations - Write down insights from the analysis conducted - Provide actionable business recommendations	4
Presentation / Notebook - Overall Quality - Structure and flow - Crispness - Visual appeal - Conclusion and Business Recommendations OR - Structure and flow - Well commented code - Conclusion and Business	8

AI Mentor

Chat with AI Mentor to get instant answers to your queries related to this assignment.

Submit Assignment

⟨ Previous   Next ⟩

🔗 **Upload file**