Job-A-Thon – May 2021

Credit Card Lead Prediction

Happy Customer Bank is a mid-sized private bank that deals in all kinds of banking products, like Savings accounts, Current accounts, investment products, credit products, among other offerings.

The bank also cross-sells products to its existing customers and to do so they use different kinds of communication like tele-calling, e-mails, recommendations on net banking, mobile banking, etc.

In this case, the Happy Customer Bank wants to cross sell its credit cards to its existing customers. The bank has identified a set of customers that are eligible for taking these credit cards.

Now, the bank is looking for your help in identifying customers that could show higher intent towards a recommended credit card, given:

* Customer details (gender, age, region etc.)
* Details of his/her relationship with the bank (Channel\_Code,Vintage, 'Avg\_Asset\_Value etc.)

**Data Dictionary**

**Train Data**

|  |  |
| --- | --- |
| Variable | Definition |
| ID | Unique Identifier for a row |
| Gender | Gender of the Customer |
| Age | Age of the Customer (in Years) |
| Region\_Code | Code of the Region for the customers |
| Occupation | Occupation Type for the customer |
| Channel\_Code | Acquisition Channel Code for the Customer  (Encoded) |
| Vintage | Vintage for the Customer (In Months) |
| Credit\_Product | If the Customer has any active credit product (Home loan,  Personal loan, Credit Card etc.) |
| Avg\_Account\_Balance | Average Account Balance for the Customer in last 12 Months |
| Is\_Active | If the Customer is Active in last 3 Months |
| Is\_Lead(Target) | If the Customer is interested for the Credit Card  0 : Customer is not interested  1 : Customer is interested |

**Test Data**

|  |  |
| --- | --- |
| Variable | Definition |
| ID | Unique Identifier for a row |
| Gender | Gender of the Customer |
| Age | Age of the Customer (in Years) |
| Region\_Code | Code of the Region for the customers |
| Occupation | Occupation Type for the customer |
| Channel\_Code | Acquisition Channel Code for the Customer  (Encoded) |
| Vintage | Vintage for the Customer (In Months) |
| Credit\_Product | If the Customer has any active credit product (Home loan,  Personal loan, Credit Card etc.) |
| Avg\_Account\_Balance | Average Account Balance for the Customer in last 12 Months |
| Is\_Active | If the Customer is Active in last 3 Months |
|  |  |

**Sample Submission**

This file contains the exact submission format for the predictions. Please submit CSV file only.

|  |  |
| --- | --- |
| **Variable** | **Definition** |
| ID | Unique Identifier for a row |
| Is\_Lead | (Target) Probability of Customer showing interest (class 1) |

How to Make a Submission?

* All Submissions are to be done at the solution checker tab.
* For a step by step view on how to make a submission check the below video

Evaluation

The evaluation metric for this competition is [roc\_auc\_score](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_auc_score.html)across all entries in the test set.

**Public and Private Split**

Test data is further divided into Public 30% and Private 70%

* Your initial responses will be checked and scored on the Public data.
* The final rankings would be based on your private score which will be published once the competition is over.

**Guidelines for Final Submission**

Please ensure that your final submission includes the following:

1. Solution file containing the predicted probabilities of Response for the customers
2. A Zipped file containing code & approach (**Note that both code and approach document are mandatory for shortlisting**)

* **Code:** Clean code with comments on each part
* **Approach:**Please share your approach to solve the problem (doc/ppt/pdf format). It should cover the following topics:
  1. A brief on the approach, which you have used to solve the problem.
  2. What data-preprocessing / feature engineering ideas really worked? How did you discover them?
  3. What does your final model look like? How did you reach it?

**Hackathon Rules**

1. *The final standings would be based on private leaderboard score.*
2. Setting the final submission is recommended. Without a final submission, the submission corresponding to best public score will be taken as the final submission
3. Use of external data is prohibited
4. Use of ID variable is not allowed as part of the model
5. You can only make **10 submissions** per day
6. Entries submitted after the contest is closed, will not be considered
7. The code file pertaining to your final submission is mandatory while setting final submission
8. Throughout the hackathon, you are expected to respect fellow hackers and act with high integrity.
9. Analytics Vidhya holds the right to disqualify any participant at any stage of the competition if the participant(s) are deemed to be acting fraudulently.
10. Use of multiple Login IDs will lead to immediate disqualification