JOB-A-THON

Online 26-02-2021 12:00 AM to 28-02-2021 11:59 PM

9702 Registered

Job Opportunities with Top Companies Prizes

ENDS IN

1 DAYS 7 HOURS 27 MIN

Registered

About Problem Statement Solution Checker My Submissions Leaderboard Discuss Status

Health Insurance Lead Prediction

Your Client FinMan is a financial services company that provides various financial services like loan, investment funds, insurance etc. to its customers. FinMan wishes to cross-sell health insurance to the existing customers who may or may not hold insurance policies with the company. The company recommend health insurance to it's customers based on their profile once these customers land on the website. Customers might browse the recommended health insurance policy and consequently fill up a form to apply. **When these customers fill-up the form, their Response towards the policy is considered positive and they are classified as a lead.**

Once these leads are acquired, the sales advisors approach them to convert and thus the company can sell proposed health insurance to these leads in a more efficient manner.

Now the company needs your help in building a model to predict whether the person will be interested in their proposed Health plan/policy given the information about:

- Demographics (city, age, region etc.)
- Information regarding holding policies of the customer
- Recommended Policy Information

Data Dictionary

Train Data

Variable	Definition
ID	Unique Identifier for a row
City_Code	Code for the City of the customers
Region_Code	Code for the Region of the customers
Accomodation_Type	Customer Owns or Rents the house
Reco_Insurance_Type	Joint or Individual type for the recommended insurance

Upper_Age	Maximum age of the customer
Lower _Age	Minimum age of the customer
Is_Spouse	If the customers are married to each other (in case of joint insurance)
Health_Indicator	Encoded values for health of the customer
Holding_Policy_Duration	Duration (in years) of holding policy (a policy that customer has already subscribed to with the company)
Holding_Policy_Type	Type of holding policy
Reco_Policy_Cat	Encoded value for recommended health insurance
Reco_Policy_Premium	Annual Premium (INR) for the recommended health insurance
Response (Target)	0 : Customer did not show interest in the recommended policy, 1 : Customer showed interest in the recommended policy

Test Data

Variable	Definition
ID	Unique Identifier for a row
City_Code	Code for the City of the customers
Region_Code	Code for the Region of the customers
Accomodation_Type	Customer Owns or Rents the house
Reco_Insurance_Type	Joint or Individual type for the recommended insurance
Upper_Age	Maximum age of the customer
Lower _Age	Minimum age of the customer
Is_Spouse	If the customers are married to each other (in case of joint insurance)
Health_Indicator	Encoded values for health of the customer
Holding_Policy_Duration	Duration (in years) of holding policy (a policy that customer has already subscribed to with the company)
Holding_Policy_Type	Type of holding policy
Reco_Policy_Cat	Encoded value for recommended health insurance
Reco_Policy_Premium	Annual Premium (INR) for the recommended health insurance

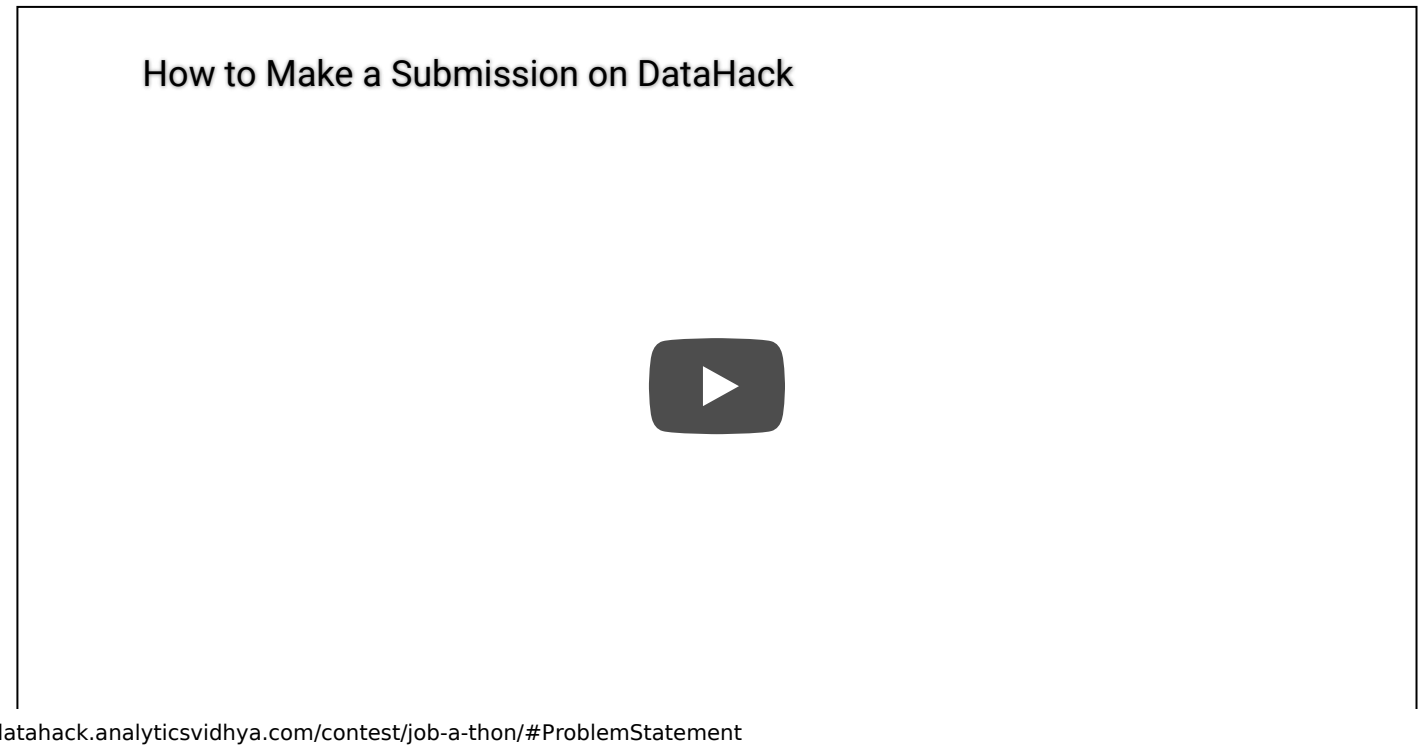
Sample Submission

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

Variable	Definition
ID	Unique Identifier for a row
Response	(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 across all entries in the test set.

Public and Private Split

Test data is further divided into Public 40% and Private 60%

- 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.

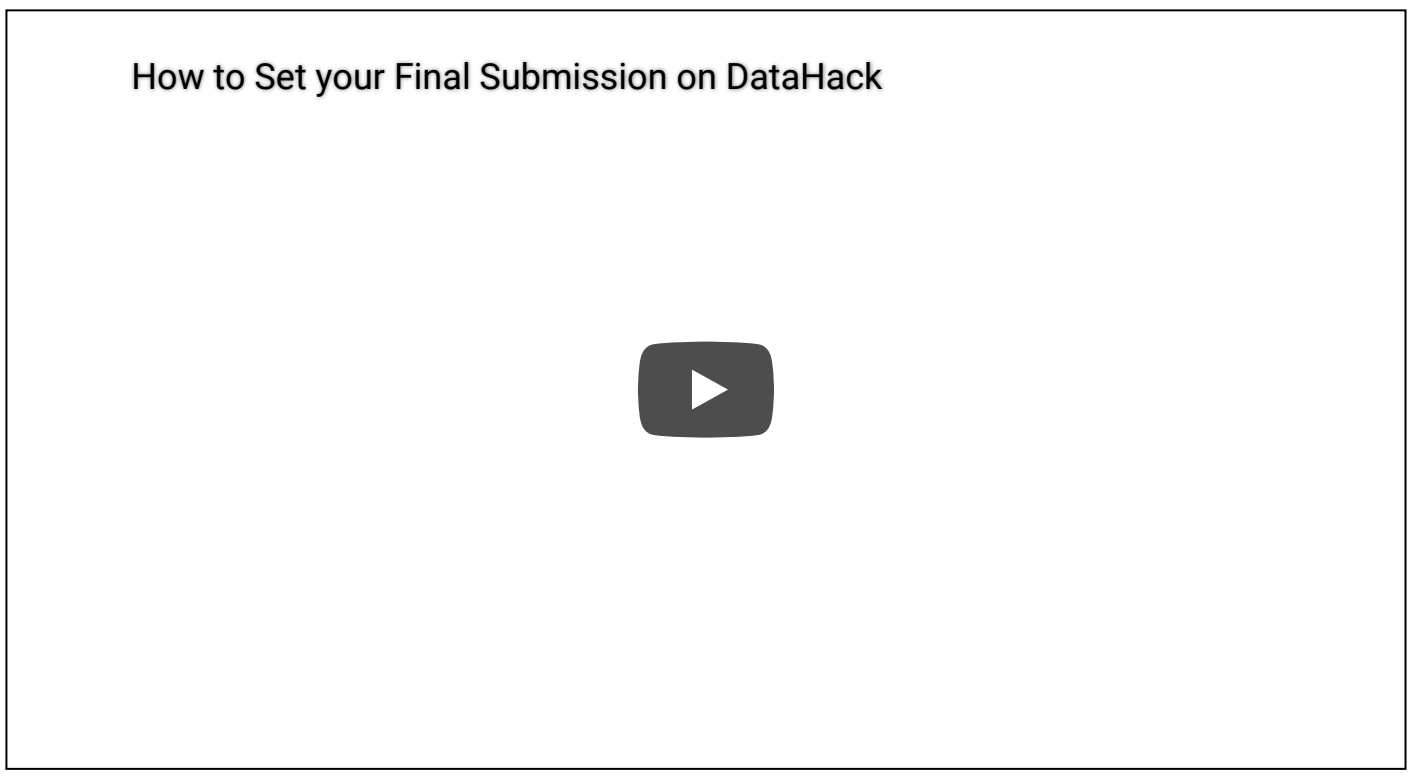
Guidelines for Final Submission

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?

How to Set Final Submission?



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

Data

Test File | Train File | Sample Submissions