08 : 22 : 27 : 03
DAY HRS MIN SEC

To all the money we've spent ...h Machine Learning Challenge

LIVE

Mar 31, 2021, 10:00 PM IST - Apr 30, 2021, 10:00 PM IST

Reduce marketing waste

Max. score: 100

You want to reduce marketing waste and aim your marketing initiatives only at those customers who will benefit from your product. This will result in the following:

- Increased business
- New customers who are compatible with your organization
- Seamless transactions with a higher success rate
- More profit with fewer obstacles

Task

Your company has products that can be used for hiring assessments. Your task is to predict the probability percentage that a client will purchase a product from the features provided in the dataset that is given.

Dataset description

The dataset folder contains the following files:

- **train.csv**: 7007 x 23
- **test.csv**: 2093 x 22
- **sample_submission.csv**: 5 x 2

The columns provided in the dataset are as follows:

Column name	Description
Deal_title	Represents a unique title for each deal
Lead_name	Represents the name of a lead
Industry	Represents the industry that a lead belongs to
Deal_value	Represents the value of a deal between a lead and

	your company (in Dollars)
Weighted_amount	Represents a value that is estimated revenue times a probability
Date_of_creation	Represents the date when a deal's pipeline was created
Pitch	Represents the different types of products that your company offers to a lead
Contact_no	Represents the contact details of a lead (masked)
Lead_revenue	Represents the lead company's revenue (in Dollars)
Fund_category	Represents the type of funding that a lead possesses
Geography	Represents the geographical location of a lead (country)
Location	Represents the geographical location of a lead (state or city)
POC_name	Represents the lead's point of contact's name
Designation	Represents the lead POC's designation
Lead_POC_email	Represents the lead POC's email address
Hiring_candidate_role	Represents the job role that a lead wants to hire
Lead_source	Represents the source from which the lead is generated
Level_of_meeting	Represents the level of a meeting with the lead. • Level 1: Introductory call • Level 2: Demo call • Level 3: Pre-sales call
Last_lead_update	Represents the communication update between a lead and your company
Internal_POC	Represents the name of the employee who has generated the lead
Resource	Represents whether your company has enough resources to satisfy a lead's requirements
Internal_rating	Represents a rating (1-5) given to a lead
Success_probability	Represents the probability that a lead will buy a product or onboard

Evaluation metric

```
score = max(0, 100-np.sqrt(metrics.mean_squared_error(actual, predicted)))
```

Result submission guidelines

- The index is **Deal_title** and the target is the **Success_probability** column.
- The submission file must be submitted in **.csv** format only.
- The size of this submission file must be 2093 x 2.

Note: Ensure that your submission file contains the following:

- Correct index values as per the test file
- Correct names of columns as provided in the **sample_submission.csv** file

Download dataset

Upload Prediction File

Please upload the prediction file in the format as stated in the problem.

Choose File No file chosen

Submit & Evaluate

Upload Source Files

You need to submit a zip or tar archive consisting of a text file explaining your approach, details about feature engineering, tools you used and the relevant source files.

Choose File No file chosen

Upload

Developers Blog

Developers Wiki

Competitive Programming

Start a Programming Club

Practice Machine Learning

?