STAGE

Machine Learning Challenge 2 - Round 1

Save & Exit

00D : 12H : 26M : 46S

Questions

1. Machine Learning Challenge . Attempted **()**

Attempts Remaining

2 / 15

Attempts made

Attempt 2	0.0000
Attempt 1	0.0000

View Leaderboard

Note:

- You can do multiple submissions.
- Your highest score will be considered

Machine Learning Challenge

Help

The data given is of cash deposits made in particular branches of a bank. Please go through the descriptions variables as below for conceptional understanding of the data.

- **Serial Number**: Unique identifier for each data point relationship.
- **Main Office**: Is the branch where the deposit was made the main office of the bank. 0 signifies no and 1 signifies yes.
- **Branch Number**: Internally used identification number of the branch or certain branches.
- **Establishment Date and Acquired Date**: The date at which the bank branch was established and acquired by the specific bank franchise.
- **CIty, Country and State:** Geographical Details of the bank branch.
- **Year Deposits Variables**(2013 Deposits, 2014 Deposits...): Deposit amount in U.S. dollars for specific branch for a specific year.

Objective Of The Problem: The objective is to predict the cash deposit amount for branches for the year of 2016 ("Deposits 2016" variable) against each "Serial Number" Variable of the test data only. Please note that the training data file is to be used for data modelling only and no predictions are to be made on the test data. The predictions are to be made as 3 clusters in the order as stated below:

- **Cluster A:** Value between 20,000 to 90,000 (Including 20,000 excluding 90,000)
- **Cluster B:** Values less than 20,000 (Excluding 20,000)
- **Cluster C:** Values greater than 90,000. (Including 90,000)

STAGE

Machine Learning Challenge 2 - Round 1

Save & Exit

The clusters must be present as continuous sets of predicted values in the above given order. Please view the sample submissions csv file for understanding of how the correct submission file format should be.

00D : 12H : 26M : 46S

Submission Limit: Submission limit for each team is 15. Please note that individual submission numbers would be shown to individual participants but if any other participant of the same team makes a submission then the team count would increase but would not be displayed on individual screens.

Evaluation Metric: The evaluation metric for this challenge is accuracy matching. Higher accuracy would indicate a better score and the same would be reflected on the leader board. Please note accuracy matches each value of the submission file with the correct labels, values predicted near to the actual correct value would not add to the accuracy. If the predicted and clustered file is exactly in order of the correct predictions then the score of 100 would be provided. Both order and prediction must be preserved

Please note that the problem is to be solved at local machines of participants only and the prediction file is to be uploaded as the solution. Submission of code is mandatory for all teams.

Evaluation Algorithm

Accuracy

Training Set(https://cdn.skillenza.com/files/ff074af7-2689-4879-afb2-b34cba1f59bd/ml2_train.csv)

Test Set(https://cdn.skillenza.com/files/46a0adf7-8fa4-40c6-8756-2344966e84e8/ml2_test.csv)

Sample Submission (https://cdn.skillenza.com/files/43aff11a-08db-4e72-bc61-04d302c28ab8/ml2_Sample_Submissions.csv)

out1.csv Upload file Submit

Help

Result ()

data rows mismatch