**Passing Criterion for this project:**

**Part 1 score should be grater than 7**

**Part 2 score should be greater than 0.84**

**Note : Part 2 passing score is different for each project**

Part 1

Part 1 of the project contains a quiz which is present in the section 'submit your part 1 solution here'. There are 10 questions of 1 mark each and you need to score atleast 7 out of 10 in order to pass the quiz i.e. to pass part 1 of the project.

Part 2

Why are our best and most experienced employees leaving prematurely? Use this database and try to predict which valuable employees will leave next. Fields in the dataset include:

Employee satisfaction level

Last evaluation

Number of projects

Average monthly hours

Time spent at the company

Whether they have had a work accident

Whether they have had a promotion in the last 5 years

Department

Salary

Whether the employee has left

We have given you two datasets , hr\_train.csv and hr\_test.csv . You need to use data hr\_train to build predictive model for response variable ‘left’. hr\_test data contains all other factors except “left”, you need to predict that using the model that you developed and submit your predicted values in a csv files.

You have to submit the probability scores, not the hard classes.

If you are using decision trees or random forest here, probability scores can be calculated as:

score=predict(rf\_model,newdata= testdata, type="prob")[,1]

score=predict(tree\_model,newdata= testdata, type=‘vector’)[,1]

Evaluation Criterion : auc score on test data. larger auc score, better Model

Please read through the points given below before you begin :

1. Your auc score for test data should come out to be more than 0.84

2. You are NOT required to submit R script. However in some cases , we might ask you to send your script separately in order to verify that your submissions is a result of models that you built .

3. you can submit as many times as you want, we'll put best score obtained on the respective leader board.

4. Your predictions should not contain any NA values.

5. You are are free to use any predictive modelling technique

Feel free to take to QA forum for this specific project for any clarifications. you will find this under the same tab