Problem statement:

The given dataset consists of a bank marketing campaign. The aim of the campaign was to persuade clients to create a term deposit with the IITHAML Bank. You are given with clients' data who were targeted during the campaign. The goal is to predict whether those people were persuaded to create the term deposit.

The data contains fields like age, job, education, etc about the client. The label is 1 if the client has created the term deposit or 0 otherwise. More details follow.

Class Labels:

- Not Subscribed: class label = 0
- Subscribed: class label = 1

File descriptions

- trainData.csv - the training dataset
- testData.csv - the test dataset

Data fields

- age - client's age (numeric)
- job - type of job (categorical: 'admin.','blue-collar','entrepreneur','housemaid','management','retired','self-employed','services','student','technician','unemployed','unknown')
- marital - marital status (categorical: 'divorced','married','single','unknown'; note: 'divorced' means divorced or widowed)
- education- (categorical:'basic.4y','basic.6y','basic.9y','high.school','illiterate','professional.course','university.degree','unknown')
- default- has credit in default? (categorical: 'no','yes','unknown')
- housing- has housing loan? (categorical: 'no','yes','unknown')
- loan- has personal loan? (categorical: 'no','yes','unknown')
- contact- contact communication type (categorical: 'cellular','telephone')
- month- last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')
- day_of_week- last contact day of the week (categorical: 'mon','tue','wed','thu','fri')
- campaign- number of contacts performed during this campaign and for this client (numeric, includes the last contact)
- pdays - number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)
- previous- number of contacts performed before this campaign and for this client (numeric)

- poutcome- outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')
- emp.var.rate- employment variation rate - quarterly indicator (numeric)
- cons.price.idx- consumer price index - monthly indicator (numeric)
- cons.conf.idx- consumer confidence index - monthly indicator (numeric)
- euribor3m- euribor 3 month rate - daily indicator (numeric)
- nr.employed- number of employees - quarterly indicator (numeric)
- Class- Output variable (desired target):has the client subscribed a term deposit? (binary: 0 or 1)

Evaluation Metric

Evaluation metric for this competition is AUC score. AUC, or Area Under Curve, is a metric for binary classification. From a random classifier you can expect as many true positives as false positives. Hence, AUC score for this case is 0.5. A score of a perfect classifier would be 1. Participants need to achieve AUC score as close to 1 as possible on test data.

Submission Format

Participants have to submit only one ".csv" (comma separated values) file. This submission file should contain two columns: Id and Class. Column "Id" represents the client Id, same as given in test dataset. Column "Class" should contain the predicted class of client by your algorithm and its value can be either 0 or 1. Value 1 represents clients who have agreed to subscribe to a term deposit and 0 represents client who did not subscribe.

The file should contain a header with the following format:

```
Id,Class
1,1
2,1
3,0
4,0
etc.
```

BEST OF LUCK!!!!