

COMPETITION RULES

ELIGIBILITY

1. All Deloitte USI and Deloitte India employees are eligible to participate in this competition. The participants also need to register for the USI Analytics Summit, 2016.

TEAM STRUCTURE

1. A team can have maximum of 3 members
2. All team members should have valid Deloitte email Ids
3. Details of all team members are to be entered at the time of registration; additions or deletion of team members are not permitted after registration
4. A participant cannot be registered with more than one team

SUBMISSION RULES

1. Each team can make a maximum of 5 submissions per day
2. Teams are required to make submissions before the 'First Submission Date'. A team cannot make their first submission post this date
3. *Till the Milestone date*, teams can make submissions similar to the sample submission file provided in the DATA section
4. *On and after the Milestone date*:
 - a. Teams will have to make two submissions that will be considered for the final private leaderboard (Refer Evaluation Sections) – these may be completely new submissions or re-submits.
 - b. Both the submissions will have to be supported by scripts/codes that can exactly reproduce the submissions. The codes are to be sent to the competition email address usianalyticssummitcompetitions@deloitte.com. The codes need to cover any cleaning of the data, training of the data, the models and final scoring.
 - c. In case of more than two submissions, the last two submissions supported with codes will be considered.

COMPETITION TIMELINE

1. Start Date: 21st March, 2016
2. First Submission Deadline: 15th April, 2016
3. Milestone Date: 15th April, 2016
4. End Date: 22nd April, 2016

DATA

1. All team/ team members are required to accept the competition rules before downloading the data. Also by downloading the data, the participants (all team members) agree to not use the data for any purposes outside the competition without permission from the organizers.
2. No external data is to be used for modeling apart from the data provided.

WINNERS

1. The winners would be decided on the basis of the private leaderboard (Refer Evaluation Section) score computed from the final two submissions that happened after the Milestone date. These submissions are to be supported by codes that can reproduce the exact submission files (Refer Submission rules). In case the codes are missing, the submission would not be considered as a valid submission.
2. Ties would be resolved on the basis of time of submission with early submissions ranked above later ones.
3. RMSE (root mean square error) scores equivalent to the baseline submission scores (check leaderboard for “Baseline1”) would not be considered while deciding the winners.

Any non-adherence to the rules or spirit of the competition will lead teams to be disqualified and removed from the leaderboard.

EVALUATION METRIC

The models are evaluated on the basis of their RMSE. RMSE is defined as the square root of average squared prediction error of all data points. Prediction error for a data point defined as the difference between value predicted by the model and the actual value observed. Mathematically,

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\text{predicted value}_i - \text{actual value}_i)^2}{N}}$$

N = total number of data points

Consider the following example,

ID	Actual	Predicted (from the submission file)
1	2.5	2.3
2	1.5	1.6
3	0.4	0.7
4	0.2	0.5
5	0.1	0.2

N = 5

$$RMSE = \sqrt{\frac{(2.3 - 2.5)^2 + (1.6 - 1.5)^2 + (0.7 - 0.4)^2 + (0.5 - 0.2)^2 + (0.2 - 0.1)^2}{N}}$$

EVALUATION CRITERIA & LEADERBOARD

Each time a submission is made, two RMSE measures are computed by the system – one of for the private leaderboard and one for the public leaderboard. THE WINNERS ARE DECIDED UPON THE RMSE SCORE ON THE PRIVATE LEADERBOARD.

Public Leaderboard

Displayed from 21st March, 2016 to 15th April, 2016

RMSE on random 78% of the test data set

Private Leaderboard

Displayed from 16th April, 2016 to 22nd April, 2016

RMSE on 100% of the test data set