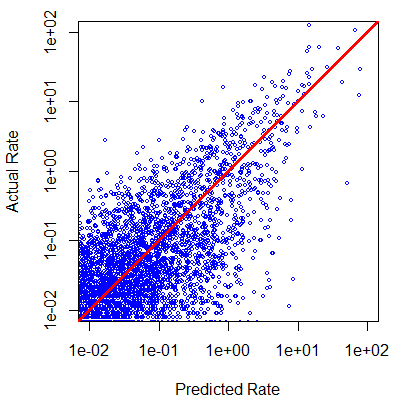
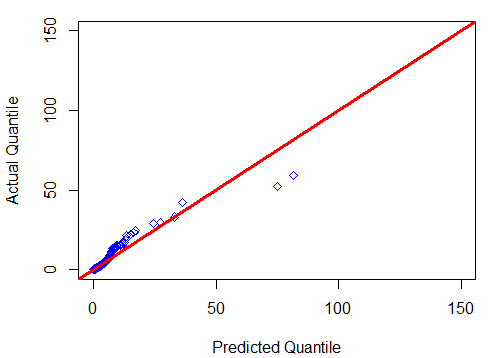
Engagement modelling for listeners on 7cot-

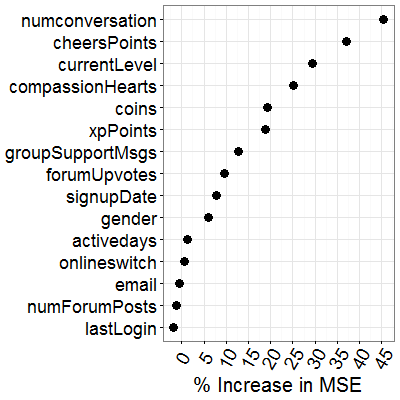
Engagement experienced from the listeners on 7cot is taken as directly proportional to the rate of messages exchanged by them with other guests and members on the website. Random forest model with n=1000 trees was taken in order to model this entity as a function of the other characteristics of the listeners as obtained using the activities performed by them on the platform. Following features were taken into account –

1. Number of coins
2. Number of group Support Msgs
3. Number of forum Upvotes
4. Number of cheers Points
5. Number of compassion Hearts
6. Number of Forum posts
7. signupDate
8. lastLogin
9. currentLevel
10. Number of xp Points
11. Number of Conversation
12. Number of Active Days
13. Gender
14. Online Switch
15. Frequency of emails

75% of the data available (randomly sampled) was used to train the model while the rest 25% of the listener’s data was used (not used for training) for testing the accuracy of the obtained model. All the listeners that registered on the website between 5th May 2013 and 16th October 2015 were taken into account. 1/3 rd of the variable were used in each tree.







Classification of users for long term engagement

A similar definition of being active and not active as in the ASONAM paper was considered here. The following features were taken into account –

1 Number of Messages from User

2 Activedays

3 AccountLogins

4 PageViewsWeb

5 ConvReq

6 PageViewApp

7 Number of Forum Posts

66% of the data was used to train the model and rest of the data was used for testing the performance of the model. The following results on the test data were obtained. 1- denoting active users and 0 denoting not active users -

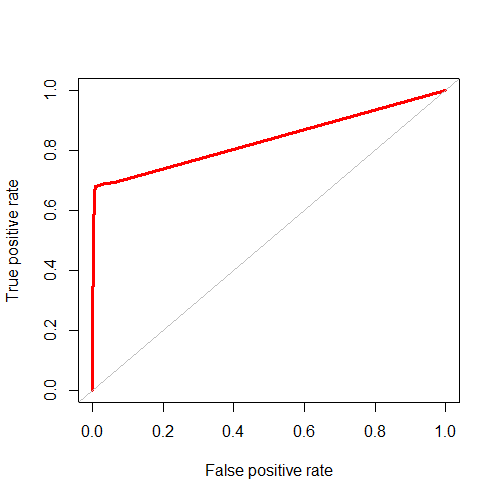
Observed

Predicted 0 1

0 9537 1113

1 67 2373

The ROC curve



Area under the ROC curve - 0.8361893

Features Importance

