Assignment 7c

Mithilesh Josyabhatla: Attempt 1

Page 1:

1 -- 2 -- 3 -- 4 -- 5 --

Question 1 (3 points)

In models `tree3` and `tree4` we examined the influence of game characteristics and category respectively on rating. Generally speaking, adding more meaningful predictors to a model will improve its predictive accuracy. Construct a regression tree that includes all the predictors used in `tree3` and `tree4`. Call this `tree5`. Specifically, `tree5` will include the following predictors: max_players, playtime, min_age, min_players, age_of_game, CardGame, Wargame, Fantasy, Fighting, Economic, Science.Fiction, Dice, Party.Game, AbstractStrategy, Childrens.Game, WorldWarII, Bluffing, Humor, Animals, Adventure, Medieval, Action.Dexterity, Deduction, Movies.TV.Radiotheme, Miniatures. What is the RMSE of `tree5`?

Question 2 (3 points)

Get H

Assignment 7c

Mithilesh Josyabhatla: Attempt 1

Page 1:

1 -- 2 -- 3 -- 4 -- 5 --

Question 2 (3 points)

The results of `tree5` may have been disheartening. Trees are flexible models that tend to overfit the data. To prevent this from occurring, rpart() has defaults that prevent the tree from becoming too complex. These defaults kept `tree5` from using all the features available. We will now change the complexity parameter to affect the size and complexity of the tree. Run tree5 again but this time explicitly set the complexity parameter to 0.005. Call this `tree6` and plot the tree. How many leaves does `tree6` have? (After you answer this question, you may also want to check the number of leaves in `tree5`.)

- 13
- 15
- 7
- 11
- 9

Assignment 7c

Mithilesh Josyabhatla: Attempt 1

Page 1:

1 -- | 2 -- | 3 --

4 -- | 5 --

Question 3 (3 points)

What is the RMSE of `tree6`?

Question 4 (3 points)

Now, keep the same predictors that you used in tree6, construct three regression trees, tree7, tree8, and tree9 with cp values of 0.001, 0.0001, and 0.1 respectively. Which of the the following regression trees has the lowest RMSE?

- tree5
- tree7
- tree9
- tree6

Assignment 7c

Mithilesh Josyabhatla: Attempt 1

- ◯ tree9
- ◯ tree6
- ◯ tree8

Question 5 (3 points)

The ultimate test of a model comes from evaluating it on a dataset not used for estimating the model. Compute RMSE for tree5, tree6, tree7, tree8, and tree9 on the "test" sample. Which of the the following regression trees has the lowest RMSE?

- ◯ tree8
- ◯ tree9
- ◯ tree5
- ◯ tree7
- ◯ tree6

Submit Quiz

0 of 5 ques

Get Help