



Question 1 (3 points)

category respectively on rating. Generally speaking, adding more meaningful predictors to a model will improve its predictive accuracy. Construct a regression In models tree3 and tree4 we examined the influence of game characteristics and

min_age, min_players, age_of_game, CardGame, Wargame, Fantasy, Fighting, Economic, Science. Fiction, Dice, Party. Game, AbstractStrategy, Childrens. Game

Specifically, tree5 will include the following predictors: max_players, playtime, tree that includes all the predictors used in tree3 and tree4. Call this tree5.

WorldWarll, Bluffing, Humor, Animals, Adventure, Medieval, Action.Dexterity, Deduction, Movies.TV.Radiotheme, Miniatures. What is the RMSE of tree5?

ints)



Mithilesh Josyabhatla: Attempt 1

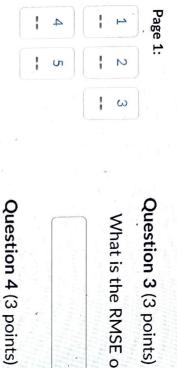


Page 1:

Cuestion c (a points)

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

Mithilesh Josyabhatla: Attempt 1



What is the RMSE of tree6?

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? Now, keep the same predictors that you used in tree6, construct three regression







tree6







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 following regression trees has the lowest RMSE?

() tree8

tree9

() tree5

tree7

() tree6