Riley Goodling

Q6

Dummy Set 1

Size: 3. Classification rate: 1.0.

Dummy Set 1 has a 100% accuracy because of attribute #5. This attribute directly determined the label for the example. Since this defining value was so strong, it allowed for a definitive result, leading to 100% accuracy.

Dummy Set 2

Size: 11. Classification rate: 0.65

Dummy Set 2 has a much lower accuracy rate than Fake Set 1. This training data set was small. There was no defining attribute that could split the data like in Dummy Set 1. The small sample size made it so that it could not find the best split to define the data, making the accuracy so low.

Car Set

Size: 408. Classification rate:0.94875.

The few attributes and many examples led this to have a high classification rate and relatively small tree size. The coverage of the data set spans over many permutations to help it classify the sets very accurately.

Connect4

Size = 41521. Classification rate: 0.7566.

The large tree and slightly low classification rate are most likely due to the fact that many combinations of the game can lead to a winner, and while there are many combinations, there is not a significant defining attribute to separate them. The tree is large because of the large amount of permutations of the board that it has to cover.

Q7

The car dataset resembles what you could see from a secondhand car dealer site. The decision tree can be used to help the user find their ideal car after putting in their most valuable attributes. The algorithm could also be implemented to help the seller decide what the best price for the car is, based on the classifications given. The algorithm can be used over and over to keep learning what customers like and eventually make better suggestions based on the factors available.

The Connect4 dataset could lead to a decision tree that acts like a value iteration agent. The agent would look at all the values on the board and try to make the best choice based on expected values of certain moves/spots. Some nodes will have a higher chance of helping the agent win later, so a token can be placed there, while others might allow the opponent to win, so a defense could be played.