

Decision Trees Problem Set

Due date: 10/13/2021

Total Points: 10

If you run out of room for an answer, use scratch paper and staple it to this sheet.

Name and Grade: _____

1. (2 points) Given the following data:

Graph the data with the horizontal axis being x_1 and the vertical axis x_2 . Shade in points that correspond to $y = 0$ and leave hollow the ones that correspond to $y = 1$. Then, draw a line that corresponds to a decision that maximizes information gain.

x_1	x_2	y
1	0	0
2	0	0
3	0	0
2	3	0
1	1	1
2	1	1
3	2	1



2. (4 points) Calculate the information gain for the above decision using the Gini Impurity Measure.

3. (2 points) Draw out a 2 layer deep decision tree using the data in Question 1 that classifies every point correctly. Specify the decisions in terms of an inequality.

4. (2 points) What is the maximum and minimum possible information gain of a decision? (this is not related to to the data)