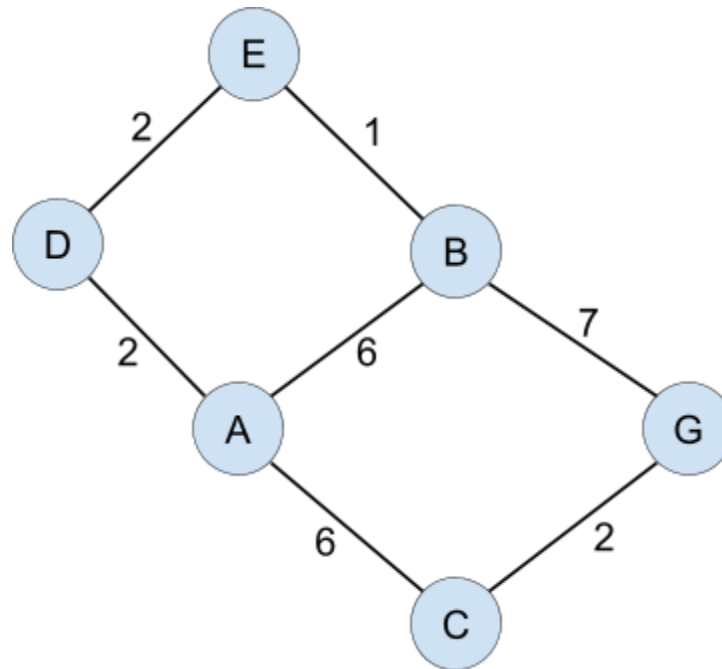Name:

NetID:

0.) Submission of practice packet (5 pts.)

1.) Given the graph below, answer the following questions (16 pts. total)

a.) Run through Dijkstra's Algorithm, starting from A and heading to G (7 pts.)

1.) A -> B: ~~2~~ 6

2.) A -> C: 6

3.) A -> D: 2

4.) A→D→E: 4

5.) A→D→E→B: 5

6.) A→D→E→B→G: 12

7.) A→B→G: 13

8.) A→C→G: 8

b.) Give a set of heuristic values for the nodes that are admissible and consistent (5 pts.)

Node	Heuristic distance from node to G
A	7
B	6
C	1
D	9
E	7
G	0

Many possible answers here

c.) Using your heuristic values, run through A* in the same manner as you did Dijkstra's (4 pts.)

$A \rightarrow B: 6+6 : 12$

$A \rightarrow C: 6+1 : 7$

$A \rightarrow D: 2+9 : 11$

$A \rightarrow C \rightarrow G: \boxed{8}$

2.) The backtracking algorithm for CSPs is a version of DFS. Why would we prefer using DFS over BFS for a CSP problem? (2 pts.)

The solutions are at the leaf nodes and it's more memory efficient

3.) Given the Minimax tree below, indicate where pruning would occur if you use alpha/beta pruning. **Note:** No need to actually show alpha and beta, but fill in every node with the right value (8 pts.)
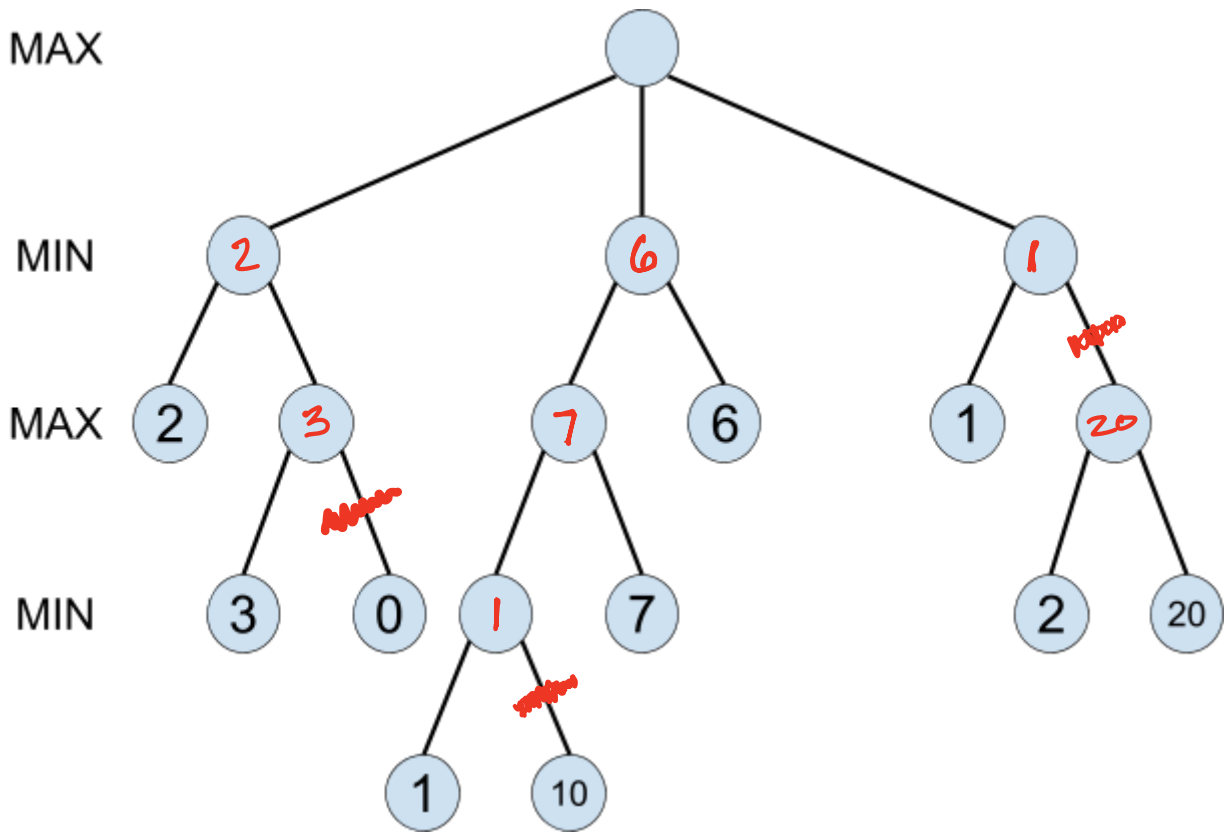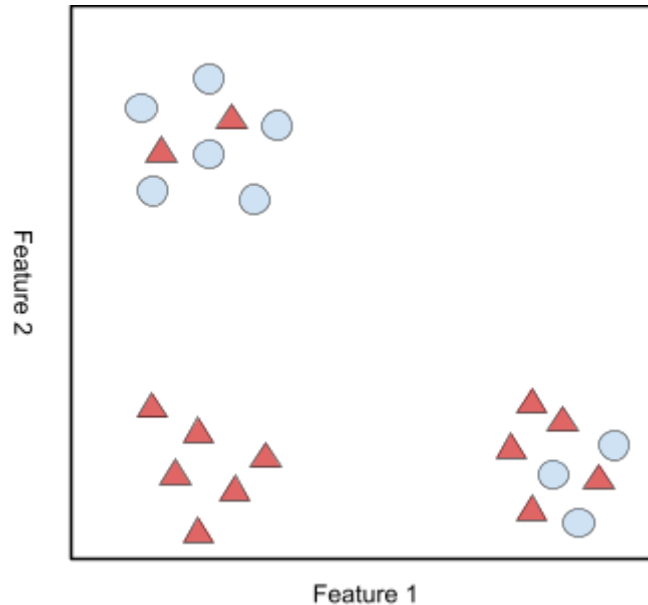

4.) What is a real-life feature that a Gaussian distribution would be a poor choice for? (2 pts.)

Net worth

5.) What is the purpose of information gain in decision trees? (1 pt.)

To help us decide which feature to split on and "where"

6.) Given the plot below, answer the following questions. (10 pts. total)


Feature 2

Feature 1

a.) What kind of task do you think this is? Is it supervised or unsupervised? (2 pts.)

Supervised classification task

b.) If we were to use Naive Bayes, should we use the Gaussian Distribution? Why or why not? (2 pts.)

No, the distribution of points looks bimodal not gaussian

c.) On the two plots below, draw a decision tree that you think has the appropriate number of splits (left) and one that is overfit (right). (2 pts.)

Straight lines, single axis
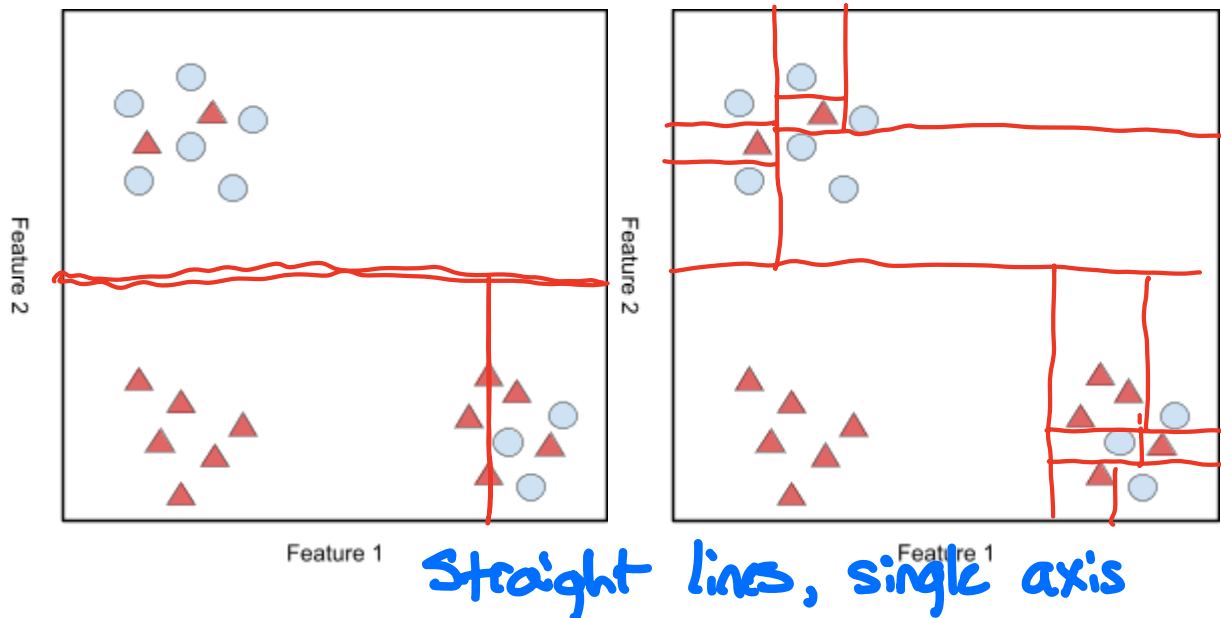
d.) What are two options decision trees have for leaf nodes that are heterogenous? (2 pts.)

Max voting or return a probability

e.) Why are decision trees considered highly explainable? (2 pts.)

they build a flowchart showing exactly where and how they make decisions

7.) Given the dataset below, answer the following questions. (8 pts. total)

X_1	X_2	Y
1	1	9
2	2	25
3	3	50
4	4	84
5	5	127

a.) Would a purely linear model exhibit high bias here or high variance? Would it be overfit or underfit? Explain your reasoning. (4 pts.)

High bias

underfit

the relationship is clearly not linear

b.) Design a regression equation that would be able to approximate the dataset. (2 pts.)

$y = 2x_1^2 + 2x_2^2 + 5$ Many possible answers

↑ should be polynomial

c.) What role does the SSE play in linear regression? Why does it matter that our resulting parameter space is convex? (2 pts.)

This is our loss function and tells us how well our Model fits the data.

Convex means we have an easy closed form solution — gradient equals 0

8.) What are emissions in the context of a Hidden Markov Model? What kind of data would we need to "learn" these? (2 pts.)

They are the possible observed States for a hidden state with associated "emission probabilities"

We'd need examples of known hidden-observed pairings to derive the domain of possible emissions and the probabilities

9.) On Homework01, why do you think Dijkstra's and A* found the same path? What, if anything, differentiated the two? (2 pts.)

They're both shortest path algorithms and they found the same shortest path

A* did it in fewer checks due to the extra heuristic information

10.) On Homework01, we only implemented backtracking for the CSP. However we could have added an AC3 component. What's an example of a constraint from HW01 that AC3 could have caught? Explain your reasoning. **Note:** I don't expect you to remember exactly every clue, just give an example of one that would be "in theme" (2 pts.)

Mr. Green was found with the gas

11.) Which lecture do you think you've learned the most from so far? Why do you think that is? I really want to improve the lectures and this feedback helps! (2 pts.)

None of them :(