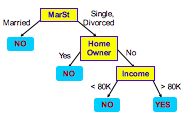
**CIS 3715 Spring 2019: Quiz 5**  STUDENT NAME: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. (1 point) When we talk about supervised learning, what is a hyperparameter? Give an example of a hyperparameter.



2. (2 point) Given a decision tree on the right, what is the decision if

the customer if Divorced, with Income of 50K and is a

Home Owner?

|  |  |  |
| --- | --- | --- |
| X1 | X2 | y |
| A | 4 | + |
| A | 0 | - |
| A | 3 | + |
| B | 2 | - |
| B | 6 | + |
| B | 1 | - |

3. (4 points) You are given the following labeled data set. What is the best first question for a decision tree? Justify your answer – the split needs to be based on some sound criterion such as misclassification error.

1. You are given a classifier that receives an example at input and produces a number at output in the range from 0 to 1. The larger the number the more confident is the classifier that the example is of positive class. The following are the outputs of the predictor and true labels on 10 test examples:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| True | + | + | + | + | + | - | - | - | - | - |
| Predicted | 0.9 | 0.3 | 0.6 | 0.8 | 0.85 | 0.4 | 0.1 | 0.85 | 0.5 | 0.55 |

We can use a threshold θ such that if the output is larger than θ the prediction is positive and if it is lower or equal the prediction is negative.

1. (2 point) What is the accuracy if the threshold is θ = 0.45
2. (1 point) Is it possible to find a threshold that results in larger accuracy?