

Chapter 8 Concept Test

Due May 13 at 12pm**Points** 1**Questions** 4**Time Limit** None

Instructions

Take this concept test after completing your pre-class preparation for chapter 8 material.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	6 minutes	1 out of 1

❗ Correct answers are hidden.

Score for this quiz: **1** out of 1

Submitted May 13 at 9:32am

This attempt took 6 minutes.

Incorrect

Question 1

0 / 0 pts

Consider a decision tree with d decisions (splits). This decision tree will split up a dataset into a certain number of chunks which correspond to the leaf nodes in the tree. How many leaf nodes will a tree with d splits have?

☐ $2d$

☐ $d-1$

☐ d

☒ 2^d

☐ d^2

☐ $d/2$ ☐ $d+1$

Since the first tree with no decisions is such that its root node is also the (only) leaf node, there is only 1 leaf node in the original tree with no splits. For each decision, a split is created which splits the leafnode into 2 leafnodes. By induction it is easy to see that there will be $d+1$ leafnodes in a d -decision tree.

Question 2

0 / 0 pts

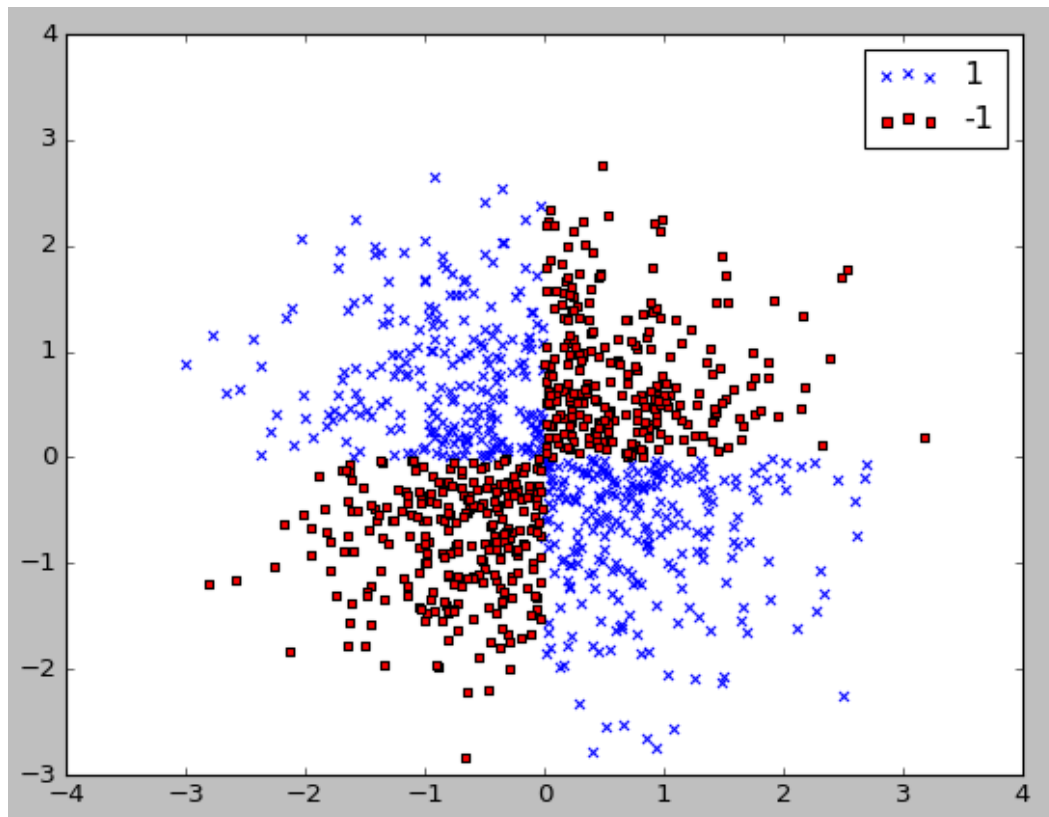
Consider a 2-dimensional dataset containing features X_1 and X_2 . If a single decision tree is run on the set, each leaf node corresponds to a region in the 2-space of the dataset. What is the maximum number of sides that each of these 2-dimensional leaf-node regions can have?

☒ 4☐ unable to determine without more information☐ 3☐ 2

The maximum number of sides for a region will occur when the region is bounded by a decision plane wherever possible. Since decision thresholds in X_1 and X_2 form orthogonal lines in the feature space, and the maximum number of bounding lines for a leaf node region is 4, the answer is 4 sides.

Question 3**0 / 0 pts**

Consider the two class dataset $\{-1, 1\}$ with 2 features as shown below. Choose the answer which best characterizes the performance of a decision tree classifier and a LDA classifier on this dataset:



- ☒ A decision tree will perform well but LDA will perform poorly
- ☐ Neither the decision tree nor LDA will perform well on this dataset
- ☐ Both the decision tree and LDA will perform well on this dataset
- ☐ A decision tree will perform poorly but LDA will perform well

A decision tree will perform well on this dataset because the region boundaries are parallel to the axes - splits will work well to separate the data.

LDA will perform poorly on this dataset because the class distributions are very distant from a Gaussian distribution and if they were modeled as Gaussian's, the class means are in the center of the dataset and the LDA-modeled distributions would likely be overlapping in the middle of the dataset - making classification difficult.

Question 4

1 / 1 pts

Please answer the following question in text form. Be specific - wherever possible, include page numbers, filenames, concept names to help your instructor understand what you are referring to: What was the most confusing aspect of the material you reviewed?

Your Answer:

The idea of boosting is still slightly unclear. Book only provides high level insight.

Quiz Score: 1 out of 1