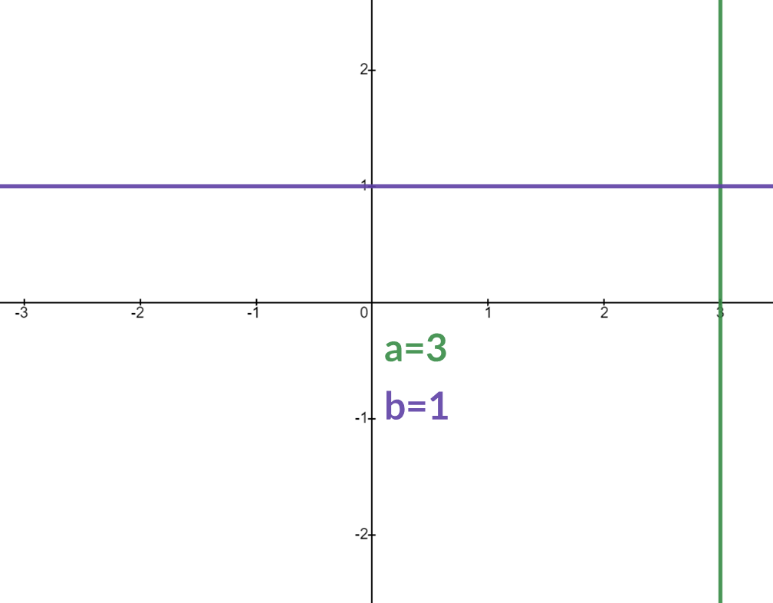
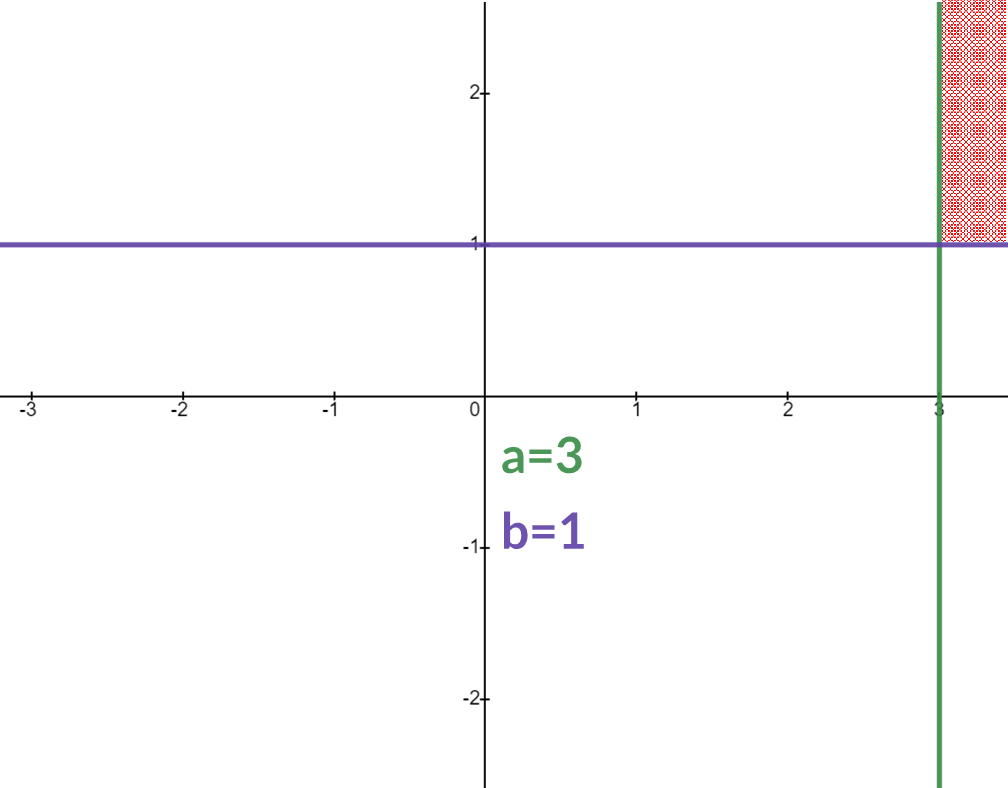
HW3 - Question 4 – Pac Learning

1. Suppose we have the classifier family: where iff . To find the VC dimension of this classifier family we need to understand it’s capabilities, we find the maximum amount of data points classifiers from can classify to every possible combination of classes. Let’s take as an example:

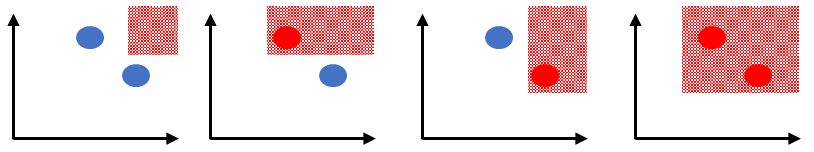


In this example the area where the classifier returns 1 is when, the area marked in red:



We will show that shatters a group of 2 samples into every possible combination of classifications but not for every group of 3 samples, hence the VC dimension would be 2.

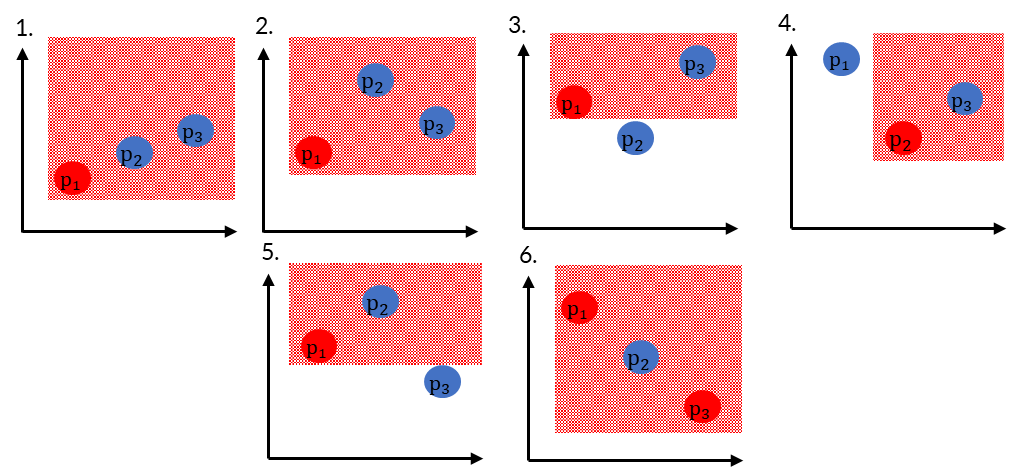
For every 2 samples that are on a straight line with a negative slope we can classify the 2 points into every possible combination of classifications using classifiers from .



The figure shows that the statement is true; The classifier is the area marked in red. The red and blue points are two types of classes.

shatters a group of 2 samples, hence the VC dimension of classifiers from is at least 2. We will now show that no 3 samples can be classified into every possible combination. We have 6 cases to consider, we will assume that :

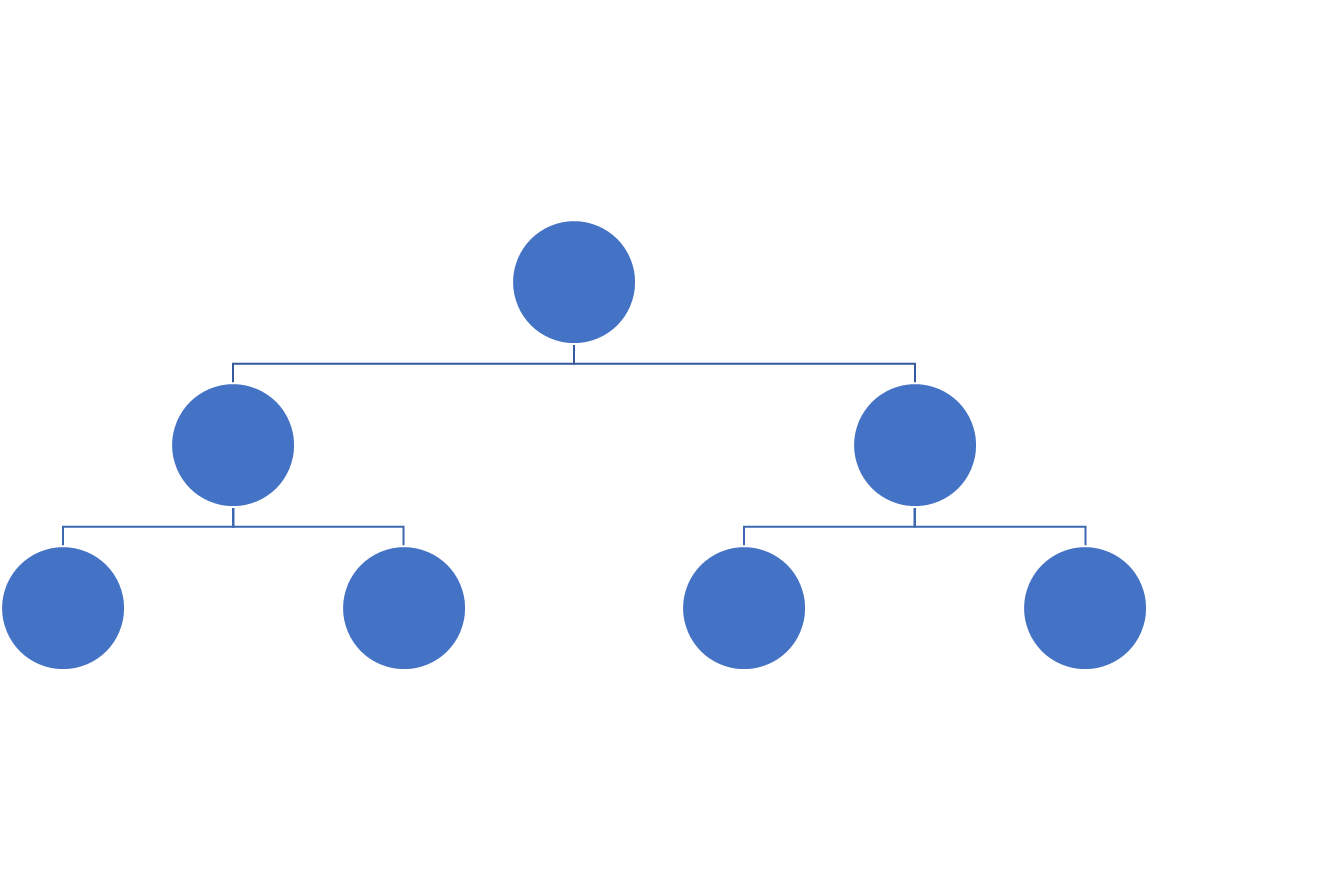
For each of these cases we will show that it is not possible to classify the points into every combination of classifications. The following scheme shows exactly that:



Each figure shows how we cannot shatter each case mentioned above.

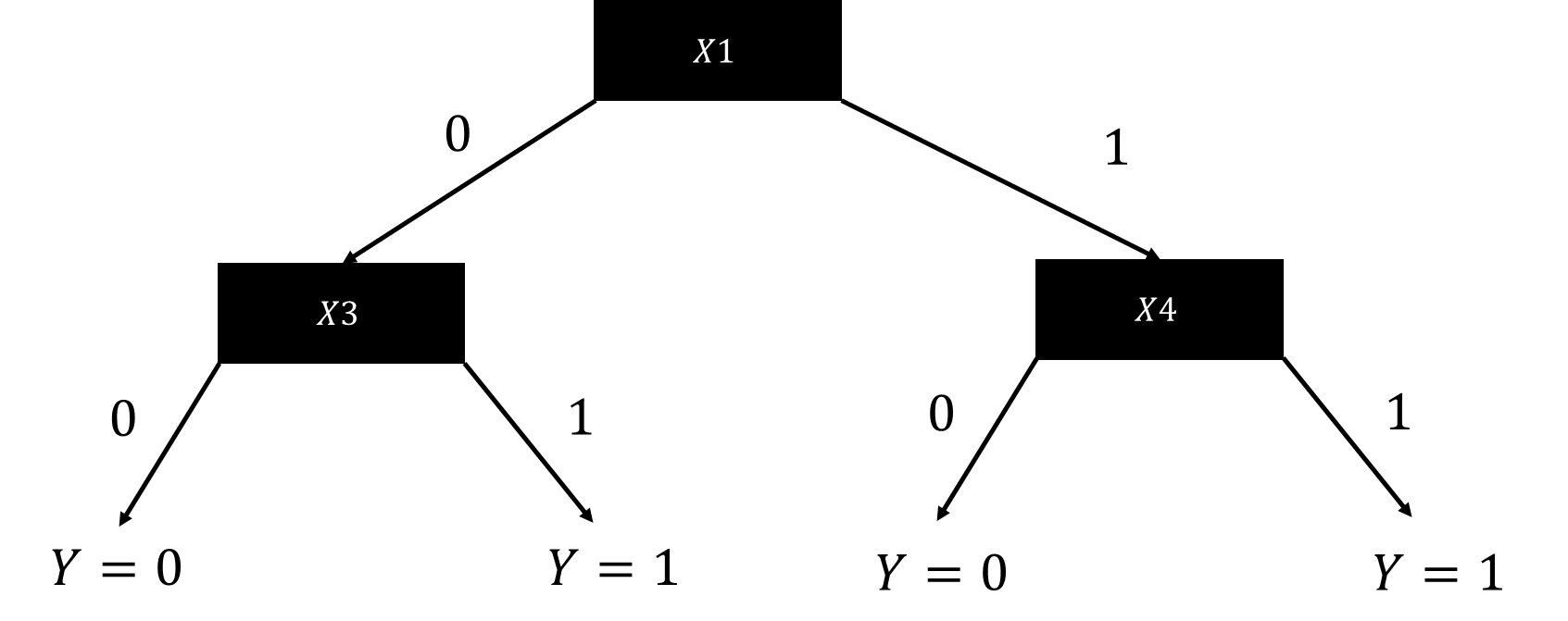
It is impossible to shatter **every** 3 points with the classifier family , hence the VC dimension of is 2. ▪

1. Boolean variables, is calculated by: , i.e is a binary classifier. We try to learn using a decision tree of depth 2. Let’s define .



A decision tree of depth 2.

Since contains a combination of decision trees, there is a finite amount of trees we can create of depth 2 and 4 leaves, so is finite. We will show that the claim of realizability applies to the problem by showing the decision tree that classifies all points correctly. The truth table for Y is the following:



This decision tree is of depth 2 and classifies each prediction precisely as intended, there isn’t any thresholds implemented but equality to 0 or 1.

Since the assumption of realizability holds, and is finite, the lower bound of the classifier family is when is the group of all decision trees of depth 2.