

Problem 1

We will use the dataset below to learn a decision tree which predicts if students pass machine learning course (True or False), based on their previous GPA (High, Medium, or Low) and whether they studied this course (True or False).

Table 1		
GPA	Studied	Passed
Low	False	False
Low	True	False
Medium	False	False
Medium	True	True
High	False	True
High	True	True

For this problem, you can write your answers using \log_2 .

- (1) What is the entropy $H(\text{Passed})$?
- (2) What is the entropy $H(\text{Passed} \mid \text{GPA})$?
- (3) What is the entropy $H(\text{Passed} \mid \text{Studied})$?

Problem 2

We have 8 examples, as shown in Table 2, which shows whether people play tennis under difference conditions: 1) Weather (Sunny or Cloudy); 2) Temperature (Hot, Mild, or Cool); 3) Wind (Weak or Strong).

Table 2			
Weather	Temperature	Wind	Play Tennis
Sunny	Hot	Weak	No
Cloudy	Mild	Strong	No
Cloudy	Hot	Weak	Yes
Sunny	Mild	Strong	Yes
Sunny	Cool	Strong	No
Cloudy	Hot	Strong	No
Sunny	Mild	Weak	Yes

Please use Naïve Bayes Algorithm to answer the following questions.

- (1) Please predict whether people will play tennis when the weather is sunny, the temperature is hot, and the wind is strong.
- (2) Please predict whether people will play tennis when the weather is sunny, the temperature is cool, and the wind is weak.

Problem 3

Recall that Adaboost learns a classifier H using a weighted sum of weak learners h_t as follows

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$$

In this question we will use decision trees as our weak learners, which classify a point as $\{1, -1\}$ based on a sequence of threshold splits on its features (here x, y).

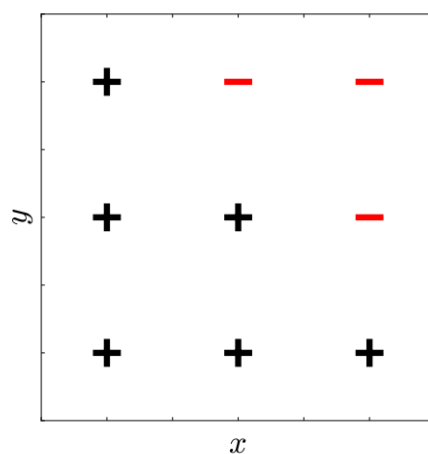


Figure 1

1. Assume that our weak learners are decision trees of depth 1 (i.e. decision stumps), which minimize the weighted training error.
 - (1) Using the dataset in Fig. 1, draw the decision boundary learned by h_1 .
 - (2) On the dataset in Fig. 1, circle the point(s) with the highest weights on the second iteration, and draw the decision boundary learned by h_2 .
 - (3) On the dataset in Fig. 1, draw the decision boundary of $H = \text{sign}(\alpha_1 h_1 + \alpha_2 h_2)$. (Hint: you do not need to explicitly compute the α 's).

2. Now assume that our weak learners are decision trees of depth 2, which minimize the weighted training error.
- (1) Using the dataset in Fig. 1, draw the decision boundary learned by h_1 .
 - (2) On the dataset Fig. 1, circle the point(s) with the highest weights on the second iteration, and draw the decision boundary learned by h_2 .
 - (3) On the dataset in Fig. 1, draw the decision boundary of $H = \text{sign}(\alpha_1 h_1 + \alpha_2 h_2)$.
(Hint: you do not need to explicitly compute the α 's).

Problem 4

We have 8 points: (3,1), (3,2), (4,1), (4,2), (1,3), (1,4), (2,3), (2,4). Please cluster it using the K-means algorithm. Let the initial cluster centers be (0,4) and (3,3) respectively. Write down the calculation process in detail. (The distance metric chooses Euclidean distance. When the previous and subsequent two clusters are consistent, stop the iteration).