Exam 2: Boosting Models

Problem 1. Adaboost. Given the data.

x_1	x_2	y
1	2	1
2	5	1
3	4	-1
4	0	-1
0	1	-1

Suppose we use this data to train an adaboost model with the learning rate $L = 1$ and obtain the three stumps as follows.

- Stump 1: $I(x_1 < 2.5)$
- Stump 2: $I(x_2 > 1.5)$
- Stump 3: $I(x_2 > 4.5)$

a. Compute the weight of the data in the first round.

b. After the first round, which observations should have the weights increased? Which observations should have the weights decreased? Explain why.

c. Compute the weights of the data in the second round and third round.

d. Compute the voting powers of the three stumps. Which stump has the highest voting power?

e. Draw the decision boundary of the adaboost.

f. Compute the error (misclassification) of the adataboost.

Problem 2. Gradient Boosting Given the data.

X	y
-1	1
2	2
4	5

We will train a gradient boosted tree (stump) on this data with the learning rate of 1.

a. Calculate the prediction of the first round.

b. Calculate the target of the second round

c. Suppose that the stump in the second round is: $X < 3$. Calculate the predictions of the second round. What do the second round predict?

d. What are the final predictions of the gradient boosting model for y in the training data after the second round?

e. Calculate the target of the third round.

f. Suppose that the stump in the third round is: $X < 0.5$. Calculate the predictions of the third round. What do the third round predict?

g. What are the final predictions of the gradient boosting model for y in the training data after the third round?