

Quiz 4, Problem 1: Adaboost (50 pts)

The Science of Deduction

You have just acquired an aging problem set about Adaboost from the year 1995. Because many of the entries are no longer legible, you decide to employ your knowledge of Adaboost to determine some of the missing information.

Below, you can see part of table that was used in the problem set. There are columns for the Round # and for the weights of the six training points (A, B, C, D, E, and F) at the start of each round. Some of the entries, marked with "?", are impossible for you to read.

Note: Questions about this table are located on the next few pages. It's depicted here just to give you a glimpse of the whole table.

	Round #	weight _A	weight _B	weight _C	weight _D	weight _E	weight _F
	1	?	?	1/6	?	?	?
①	2	?	?	?	?	?	?
	⋮						
	219	?	?	?	?	?	?
②	220	1/14	1/14	7/14	1/14	2/14	2/14
③	221	1/8	1/8	7/20	1/20	1/4	2/20
	⋮						
④	3017	1/2	1/4	1/8	1/16	1/16	0
	⋮						
⑤	6034	1/8	3/8	1/8	2/8	3/8	1/8
	⋮						
⑥	8888	?	1/10	3/10	?	?	?
	8889	3/20	1/10	?	?	?	3/20
	⋮						
⑦	9999	?	?	?	?	?	?
	10000	2/30	2/30	?	9/30	10/30	5/30

In the following problems, you may assume that non-consecutive rows are independent of each other, and that a classifier with error $< \frac{1}{2}$ was chosen at each step.

1 (8 pts) According to a scribbled note in the margin, the weak classifier chosen in **Round 1** correctly classified training points A, B, C, and E but misclassified training points D and F. What should the updated weights have been in the following round, **Round 2**?

weight _A	weight _B	weight _C	weight _D	weight _E	weight _F

Show your work for partial credit.

2 (7 pts) During **Round 219**, which of the training points (A, B, C, D, E, F) must have been misclassified, in order to produce the updated weights shown at the start of **Round 220**?

Round #	weight _A	weight _B	weight _C	weight _D	weight _E	weight _F
219	?	?	?	?	?	?
220	1/14	1/14	7/14	1/14	2/14	2/14
221	1/8	1/8	7/20	1/20	1/4	2/20

List all the points that were misclassified. If none were misclassified, write NONE instead.

3 (7 pts) During **Round 220**, which of the training points (A, B, C, D, E, F) must have been misclassified in order to produce the updated weights shown at the start of **Round 221**?

List all the points that were misclassified. If none were misclassified, write NONE instead.

4 (7 pts) Your friend John observes that the weights in **Round 3017** cannot possibly be right. Why not?

Round #	weight _A	weight _B	weight _C	weight _D	weight _E	weight _F
3017	1/2	1/4	1/8	1/16	1/16	0

5 (7 pts) Your friend Sherlock observes that the weights in **Round 6034** cannot possibly be right. Why not?

Round #	weight _A	weight _B	weight _C	weight _D	weight _E	weight _F
6034	1/8	3/8	1/8	2/8	3/8	1/8

6 (7 pts) You observe that there must also have been a mistake in the weights of **Round 8888** or **Round 8889**. Why?

Round #	weight _A	weight _B	weight _C	weight _D	weight _E	weight _F
8888	?	1/10	3/10	?	?	?
8889	3/20	1/10	?	?	?	3/20

7 (7 pts) Finally, suppose you know that at least half of the training points were classified correctly during **Round 9999**. Knowing this, which of the training points (A, B, C, D, E, F) were given a lower weight in the next round, **Round 10000**?

Round #	weight _A	weight _B	weight _C	weight _D	weight _E	weight _F
9999	?	?	?	?	?	?
10000	2/30	2/30	?	9/30	10/30	5/30

List all the points that were given a lower weight in Round 10000. If none of the points were given a lower weight, write NONE instead.