Intro To ML – HW2

students:

Dor Bank - 301380416 - dorbank@gmail.com, Avi Caciularu - 203056585 - avi.c33@gmail.com

**Theory Questions**

**Question 1:**

We will prove here a lower bound: . let be a unit vector of size d, where '1' appears in the i-th place.

We will show that the set can be shattered by . To see that, we need to show that every label assignment for is consistent with some hypothesis in . let be the assignment for the vectors . Then hypothesis c as:

for every choose .

also, for the computation of the result at the output node we will set:

We get

Please note that here we assume that . the permutation for is trivial.

**Question 2:**

In this questions we use the following lemmas proved in recitation 7:

Lemma 1:

Lemma 2:

Lemma 3:

a. Denote the function family of each node *j* in layer *i* as . In our case, since the node separators are independent of each other, we get that for each layer *i*, . Because there are d nodes at each layer, we can see that all are equal to our wanted . From Radon's theorem we get that for each , . from Lemma 2 we get that

b. C is simply composed from concatenating the family times. explicitly,

C=(L times). From Lemma 3 we get that

c. for every node we have parameters for the weight vector w, and 1 more for the bias . so in total we have d+1 parameters. we have layers with nodes, and 1 layer (the final one) with one node. combining it all together we get

.

d. Assume . First notice that . otherwise we get which is surely not true for a large enough .

Now, by taking the log of each size we get , where the second transition is true as stated above.

e. first notice that . Now, let there be a set of size that is shattered. Then . combining this with the inequality mentioned we get . from section d we get that . thus we conclude that .

**Question 2:**

a.

b.