ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 6 Published

Details Questions

☑ Show Question Details

Question 1 pts

Select all of the statements that are true below.

ıswer

☐ Consider two learning scenarios using the same learning algorithm L, and the same hypothesis space H. In the first scenario we can tolerate error rate of ϵ_1, and in the second, error rate of ϵ_2. If $\epsilon_1 < \epsilon_2$, then in the first scenario we will probably need to train L on more examples than in the second scenario.

ıswer

☐ When learning using a finite hypothesis space H, if the sample complexity of the learning algorithm is super-polynomial in the log of the size of H, then the concept class, C, is not PAC learnable.

☐ Let S be a set training examples of size m that is labeled according to a function $f \in C$. If the learning algorithm L is completely consistent with all m examples in S then C is PAC learnable.

☐ If the algorithm's time complexity is not polynomial, then the concept class, C, is not PAC learnable.

Question 1 pts

Is the class H_{disj}(all disjunctions over n variables) PAC learnable ?

ıswer

○ Yes, since $|H_{disj}| = 3^n$ and learnability depends on the log of this size, and there exists an algorithm that is consistent with an hypothesis in H_{disj}.

- Yes, since $|H_{disj}| \ = \ poly\,(n)$
- No, since it is NP hard to find the smallest disjunction that is consistent with the dataset.
- No, since $|H_{disj}| = 3^n$

Question

1 pts

We want to learn a weight vector by optimizing the Soft SVM formulation. We want to use the L2 loss for optimization:

$$\min_w \left(\frac{1}{2}w^T w \ + \ C \ \max\left(0, 1 - y_i w^T x_i\right)^2\right)$$

Assume that we have $1 > y_i w^T x_i$. If we want to use SGD to update our weight vector, which one of the following is the correct update equation?

ıswer

- $w_{t+1} \ = \ w_t \ - \ \eta w_t \ + \ 2C\eta\left(1 - y_i w_t^T x_i\right) y_i x_i$
- $w_{t+1} \ = \ w_t \ + \ \eta w_t \ - \ 2C\eta\left(1 - y_i w_t^T x_i\right) y_i x_i$
- $w_{t+1} \ = \ w_t \ - \ \eta w_t \ + \ 2\eta C y_i x_i$
- $w_{t+1} \ = \ w_t \ + \ \eta w_t \ - \ 2\eta C y_i x_i$
- $w_{t+1} \ = \ w_t \ - \ \eta \parallel w_t \parallel \ + 2\eta C y_i x_i$
- $w_{t+1} \ = \ w_t \ + \ \eta \parallel w_t \parallel \ - 2\eta C y_i x_i$

Question

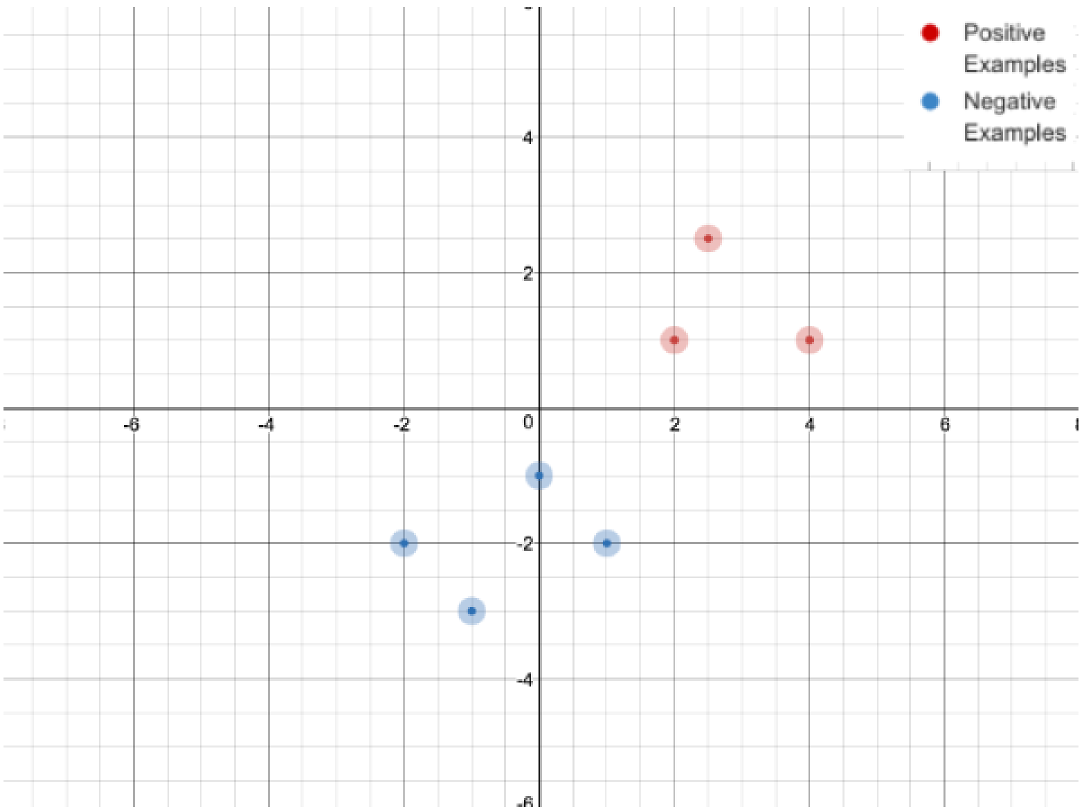
1 pts

Which weight vector will you obtain when you optimize according to the Hard SVM formulation

$$\min_{\mathbf{w}} \frac{1}{2}||\mathbf{w}||^2$$
$$\text{s.t } y^{(i)}(\mathbf{w}\cdot\mathbf{x}^{(i)} + \theta) \geq 1, \forall(\mathbf{x}^{(i)}, y^{(i)}) \in D$$

for the following data points below?

Example	x_1	x_2	y = Label
1	2	1	1
2	0	-1	-1
3	2.5	2.5	1
4	4	1	1
5	-2	-2	-1
6	-1	-3	-1
7	1	-2	-1



ıswer

- $w^T = [0.5, -0.5], \theta = 0$
- $w^T = [0.5, -0.5], \theta = -0.5$
- $w^T = [0.5, 0.5], \theta = -0.5$
- $w^T = [0.5, 0.5], \theta = 0.5$
- $w^T = [0.5, -0.5], \theta = 0.5$
- $w^T = [0.5, 0.5], \theta = 0$

Question

1 pts

Consider the hypothesis space of “origin-centered spheres”, the set
of all spheres that are centered at the origin of the three-dimensional space. The class can be formally represented as:

$$H_{\bigcirc} = \{h_r : x^2 + y^2 + z^2 = r^2 : r \in R, r > 0\}$$

For each r, h_r assigns a positive or negative label based on if the point is inside or outside of the sphere. (for example, it can assign positive a label to the point outside of the circle and vice versa.)

Is this hypothesis PAC learnable? What is the VC Dimension of the hypothesis space?

- Yes, The hypothesis space is PAC learnable. VC(H) = 3, because we can shatter a subset of size 3 but no subset of size 4 can be shattered.

ıswer

- Yes, The hypothesis space is PAC learnable. VC(H) = 2, because we can shatter a subset of size 2 but no subset of size 3 can be shattered.
- No, The hypothesis space is PAC learnable. VC(H) = 4, because we can shatter a subset of size 3 but no subset of size 4 can be shattered.
- Yes, The hypothesis space is PAC learnable. VC(H) = 3, because we can shatter a subset of size 2 but no subset of size 3 can be shattered.
- No , The hypothesis space is not PAC learnable. VC(H) = 2, because we can shatter a subset of size 2 but no subset of size 3 can be shattered.
- No , The hypothesis space is not PAC learnable. VC(H) = 3, because we can shatter a subset of size 2 but no subset of size 3 can be shattered.

Question

1 pts

We define a hypothesis space H that is the union of two intervals $[a,\ b]$, $[c,\ d]$, where $a < b < c < d$. An instance x is positive if and only if $a < x < b$ or $c < x < d$.

What is the VC dimension of H?

ıswers 4 (with margin: 0)

+ New Question + New Question Group Find Questions

Notify users this quiz has changed

Cancel Save