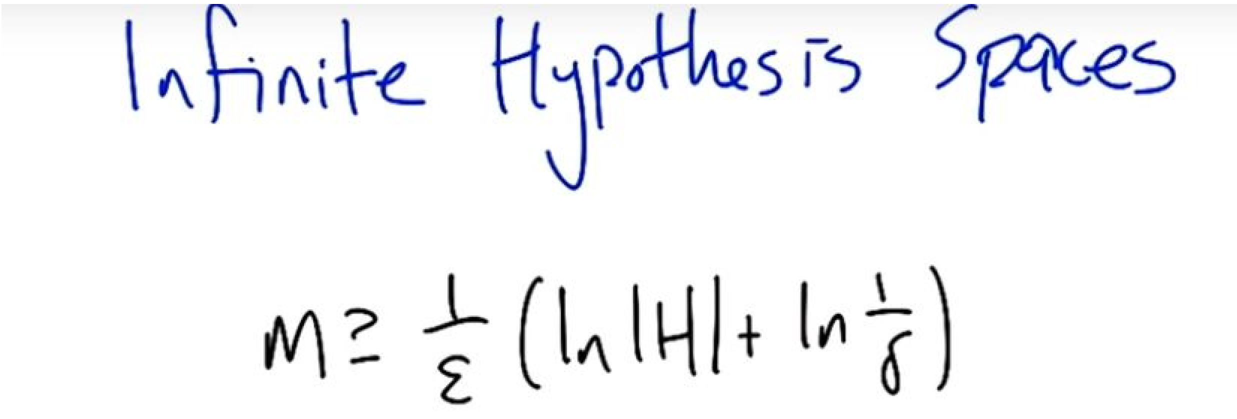
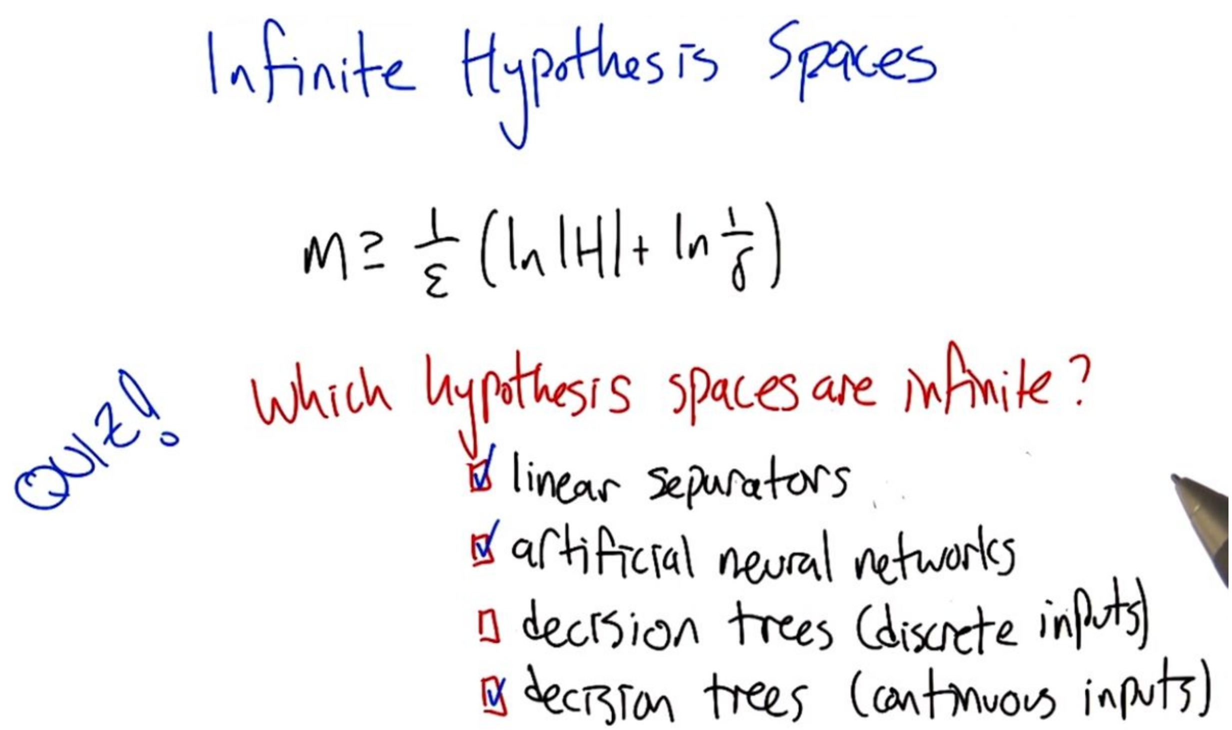
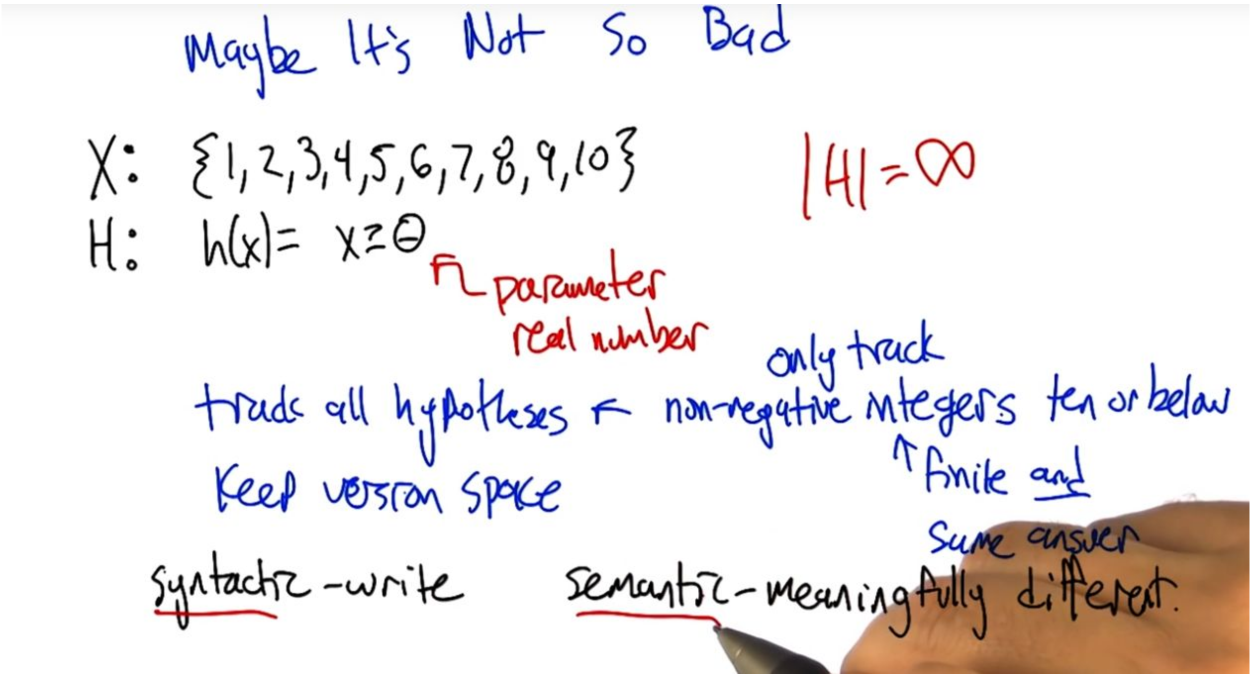
1. Infinite Hypothesis Spaces



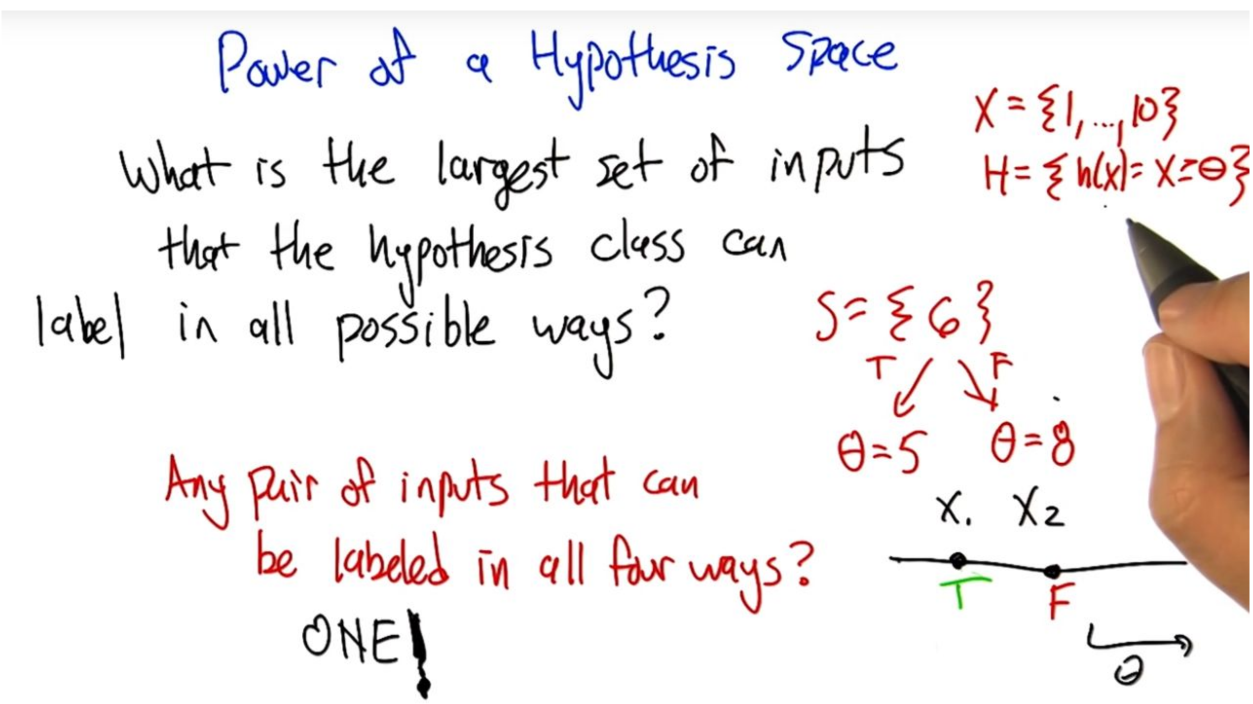
2. Quiz: Which Hypothesis Spaces



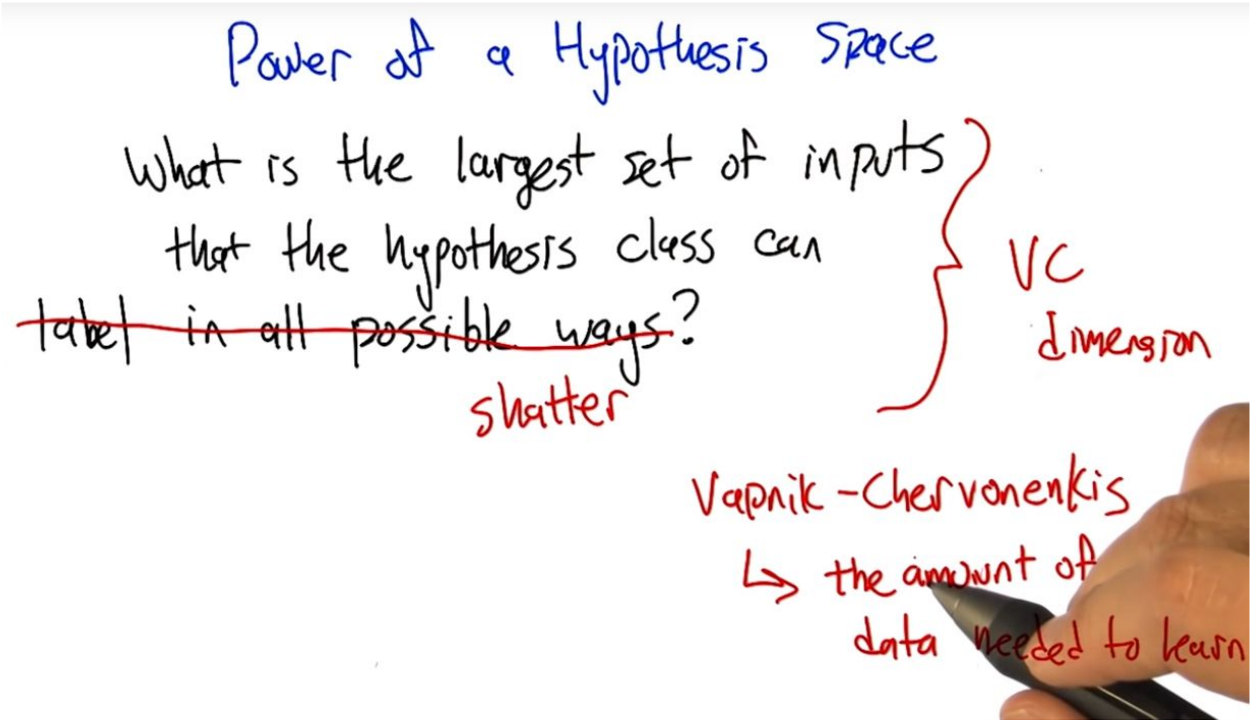
3. Maybe It Is Not So Bad



4. Power of a Hypothesis Space



5. What Does VC Stand For



6. Quiz: Internal Training

This is one of these situations where it's going to be really helpful to apply the notion of VC dimension if we think we'd like to be able to learn from a finite set of data. Which, you know, generally we like that. So how do we figure out what the VC dimension is?
We want to know, what is the largest set of inputs that we can label in all possible ways, using hypotheses from H.
Alright, so, I want you to figure that out. Figure out the size of the largest set that we can shatter, that we can label in all possible ways using these hypotheses. And then just, you know, write it as an integer in this box.

Quiz : Interval training

$X = \mathbb{R}$

$H = \{ h(x) = x \in [a,b] \}$

parameterized by $a,b \in \mathbb{R}$

$VC = \boxed{}$

ANSWER:

M: OK, so how do we figure this out?

C: Cleverly, so I, when I, when I see things like this, I just like to be methodical, so why don't we just be methodical so, I'm going to ask the question whether the vc dimension is at least one, because it's pretty easy to think about and maybe I'll get a feel for how to get the right answer that way. OK, so is the vs dimension at least one? Well, the answer is pretty clearly yes, so if you just put a dot on the number line somewhere.

You could label it positive just by picking any a less than or equal to that point and any b greater than or equal to that point. So, if, if I were like drawing parentheses or something to indicate the interval, I could just put parenthesis around the point and that will give me a plus or brackets, that would be fine. Okay, so that's that's pretty easy. And if I wanted it to be negative, I could just put both of the brackets on either side of the point, it doesn't matter, let's say to the left. Alright, that make sense Michael?

M: That's exactly what I was thinking about, yeah. Though I would've put the brackets on the right.

C: Yeah, you would. okay, so then we could see...do the same argument for, see if the VC dimension is greater than or equal to two. So if I put two points on the line, so there are only,there're four possibilities I gotta get. Plus plus, minus plus, plus minus, and minus minus. Okay, so we gotta get plus plus, plus minus, minus plus and minus minus. So, the, the first and the last one are really easy. Actually they're all easy but you can definitely do this. So, if you want to get plus plus, you just need to put brackets so that they surround the two points, that's good. If you want to get plus minus you put the left bracket in the same place and you put the right bracket just to the right of the point, yeah, and you do the same thing for minus plus and then for minus minus you put the brackets on either side of both of the points and so, since you like it to the right I'm going to put em to the left.

M: [LAUGH] Good.

C: And there you go, that was, that was pretty easy I think. Okay so next we need to figure out whether the VC dimensions at least three. So we need three dots on a line, three, distinct dots on a line. And we've got eight possibilities but Michael I don't want you to write down those, those eight possibilities because I think I see an easy way to answer the question right away.
M: Excellent
C: So, this is a lot like the last example we did with, with the theta. Except no
M: Yeah.
C: We only have two parameters. And the problem with had with the theta was that as we moved the theta over, from left to right, we lost the ability to, to, to have a, a, a positive followed by a, a negative. So I think there's a similar thing here. So, if you label those three points this way. Plus, minus, and plus. I don't, I don't think you can do that, and that's because in order to get point one and point three in the interval, you're going to have to put the brackets on both sides of them. So you're going to have to put a, a left bracket to the left of the first point and a right bracket to the right of the third point. And that's the only way to make those two plus. But then you're always going to capture the one in the middle. So you can't actually shatter three points, with this hypothesis class.
M: Now, you have to argue though, that there isn't some other way you could arrange the three points. I don't know like, I don't know, stacking them on top of each other or something.
C: You mean vertically on top of one each other?
M: Yeah.
C: Well then they wouldn't be in R, they'd be in R2.
M: Well no, just like right on top of each other.
C: Well then they're all the same point.
M: And you can't label them. Again, you have the same problem that you can't label one of them negative and the other ones positive if they're all on top of each other.
C: Right.
M: So, so there isn't, there just isn't any way to set up these three points so that you're able to assign them all possible labels.
C: Right.
M: So, good, so that gives us two as our answer here. So, by the way, I think that you said something I think that's really important. In order to prove the lower bound, in order to prove one and two, all we had to do was come up with an example where we could shatter, right?
C: Yes, that's exactly right.
M: Right, so so that's good and that's that's really nice because otherwise we're in a heap of trouble [LAUGH] if we have to show that you can shatter every single thing. We just have to show that you can shatter one thing. So, it exists. So that whole VC dimension is really a...there exists some set of points you can shatter, not you can shatter everything.
C: That's right, and what would be an example of points that you couldn't shatter yeah, a pair of points that you couldn't shatter?

M: Well, the ones on top of one another.
C: Yeah, exactly, because you wouldn't
be able to assign them different labels.
M: Right.
C: So that would be a really bad choice, and here all we need is a good choice.
M: Right. So, if you make good choices you can shatter things, which sounds more violent than I intended. Okay but, in the third case of the VC dimension, it wasn't enough to show an example that you couldn't shatter, because, then you could do the same thing as you point out, with a VC dimension of two. Instead you have to prove that no example exists. So, there does not exist or a for all not word or something.
C: For all, not.
M: [LAUGH] Exactly. So, that, that's a, that's an interesting set of set of requirements there, right? So, proving a lower bound seems easier than proving an upper bound.
C: Though it's interesting because in this case, in cases one and two, you had to show that all the different combinations were covered, whereas in this last case we just had to give one combination that couldn't possibly be covered.
M: Yeah, but it couldn't possibly be covered no matter what we did. No matter what the input arrangement was.
C: Right.
M: Yeah.
C: Whereas in the first case, I had to show all possibilities. I mean, you know, all possible labelings but only for one example of orderings or one collection of points. So just messily doing some bad predicate calculus to, nail down what you're saying. That when we say that the answer is yes, we're saying that there exists a set of points of a certain size, since that for all labelings, no matter how we, we want to label it. There is some hypothesis that works for that labeling. But to say no, we have to do the legation of that which is not exist for all exist. Which, by standard logic rules says that, that means for all points, no matter how you arrange the points, it's not the case that for all labels. There exists hypothesis which again DeMorgan's Law its not against DeMorgan's Law to to apply this idea that says that's the same as for all arrangements of points there's some labeling where there's no hypothesis that's going to work and that's exactly how you made your argument.
M: Huh, except I didn't use DeMorgan's Law and upside down a's and backwards z's. Oh you did, oh you did.

Quiz : Interval training

$X = \mathbb{R}$

$H = \{ h(x) = x \in [a,b] \}$

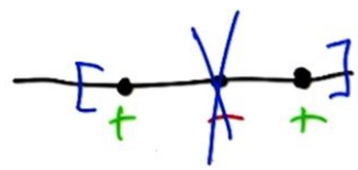
parameterized by $a,b \in \mathbb{R}$

VC= $\boxed{2}$

Is the VC dimension ≥ 1?

Yes!

Is the VC dimension ≥ 2?



Yes!

Is the VC dimension ≥ 3? No!

Quiz : Interval training

$X = \mathbb{R}$

$H = \{ h(x) = x \in [a,b] \}$

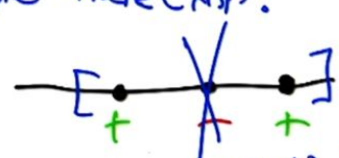
parameterized by $a,b \in \mathbb{R}$

VC= $\boxed{2}$

Is the VC dimension ≥ 1?

Yes! ← Just need <u>one</u> example : there exists!

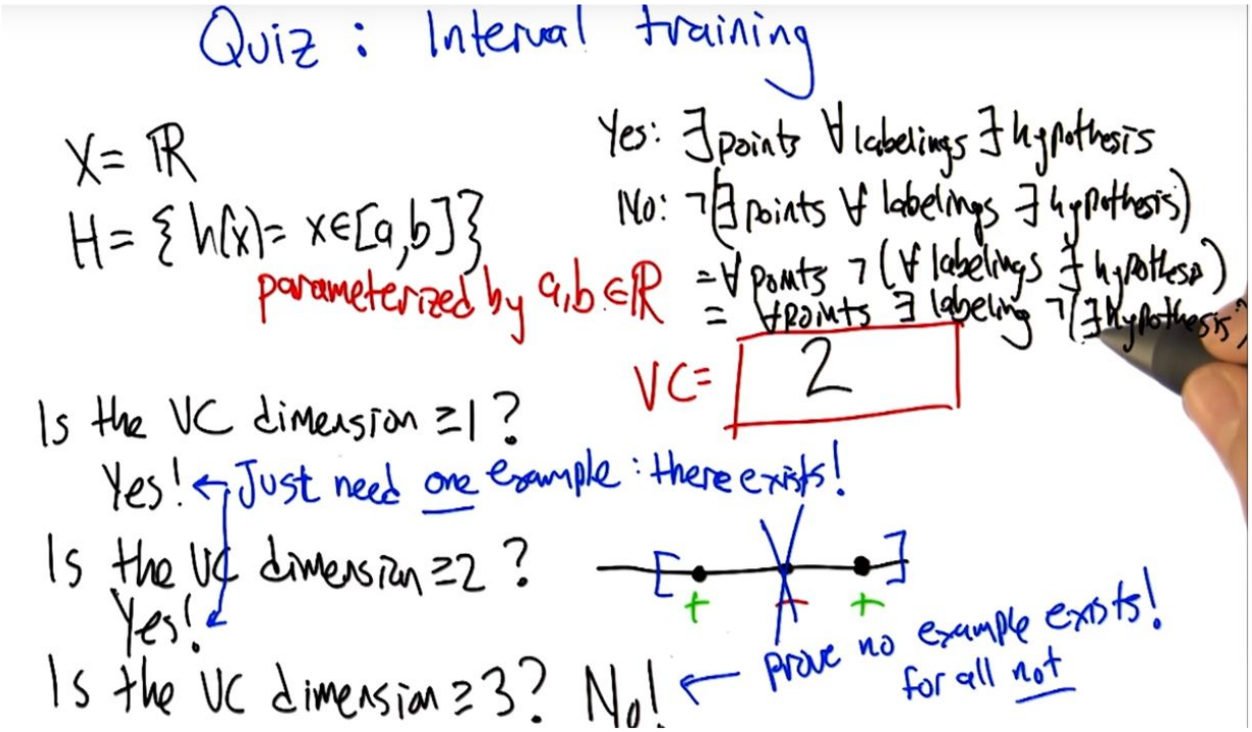
Is the VC dimension ≥ 2?



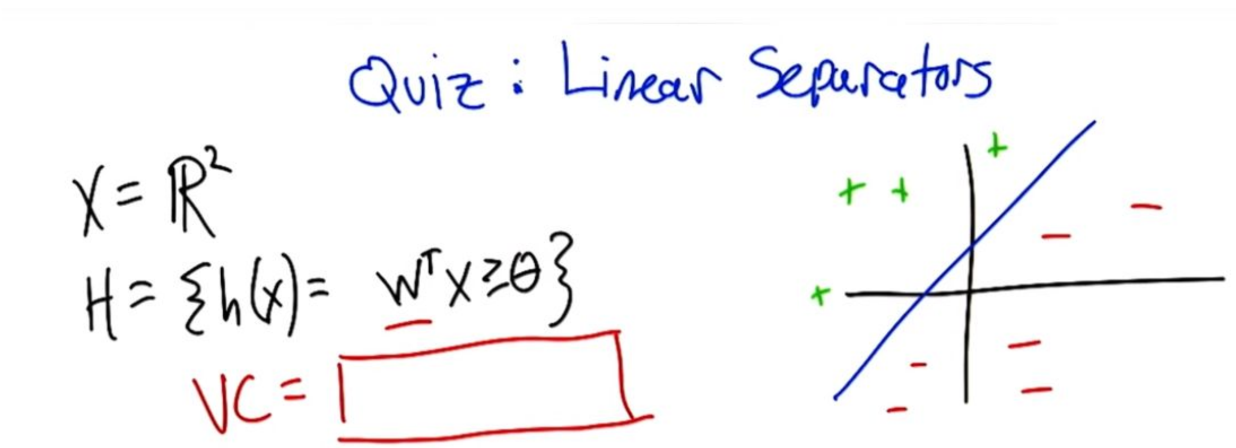
Yes!

Is the VC dimension ≥ 3? No! ← prove no example exists!

for all <u>not</u>



7. Quiz: Linear Separators

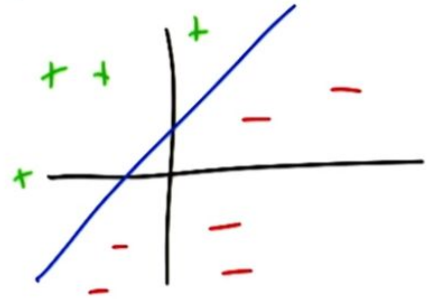


Quiz : Linear Separators

$X = \mathbb{R}^2$

$H = \{ h(x) = \underline{w}^T x \geq \theta \}$

$VC = \boxed{}$



Is $VC \geq 1$? Yes!

Is $VC \geq 2$? Yes!

Is $VC \geq 3$? Yes!

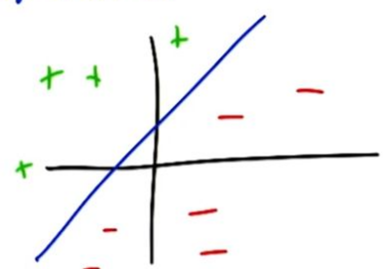
Is $VC \geq 4$? No?

Quiz : Linear Separators

$X = \mathbb{R}^2$

$H = \{ h(x) = \underline{w}^T x \geq \theta \}$

$VC = \boxed{3}$

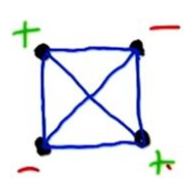


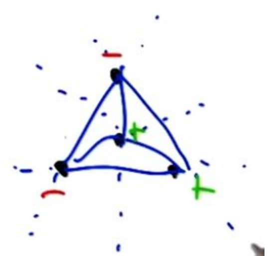
Is $VC \geq 1$? Yes!

Is $VC \geq 2$? Yes!

Is $VC \geq 3$? Yes!

Is $VC \geq 4$? No!





8. The Ring

Hypotheses

one dimension	1	θ
interval	2	a, b
two dimension	3	w, θ
three dimension	4	✓
d-dimensional hyperplane	d+1	

VC dimension often number of parameters!

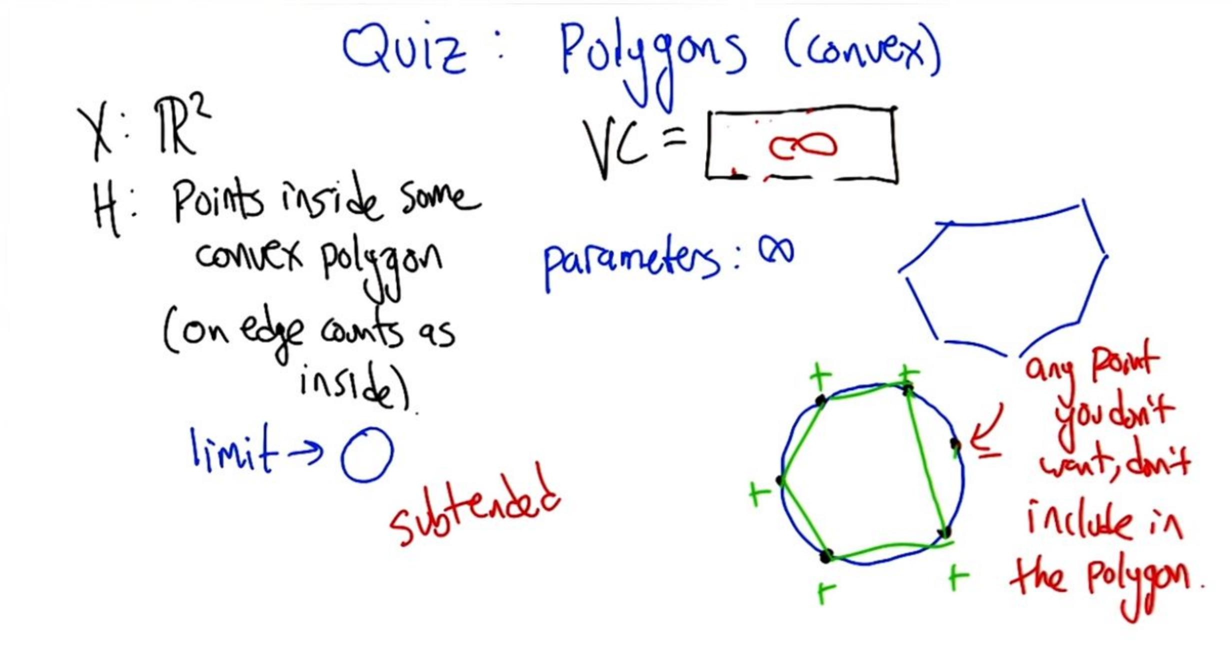
9. Quiz: Polygons

Quiz: Polygons (convex)

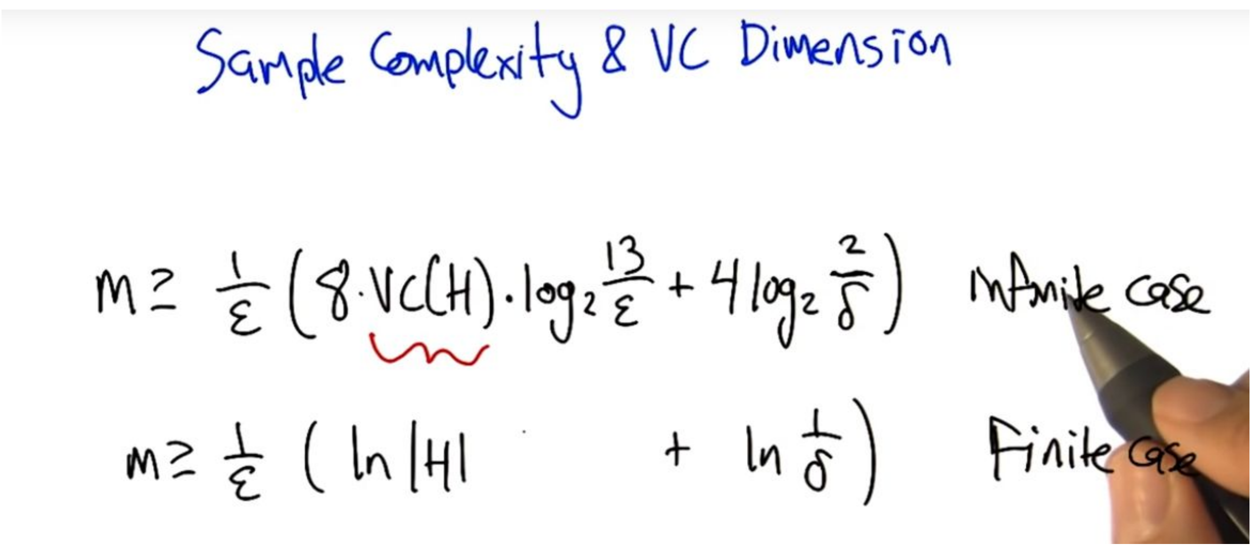
X: $\mathbb{R}^2$

H: Points inside some convex polygon (on edge counts as inside)

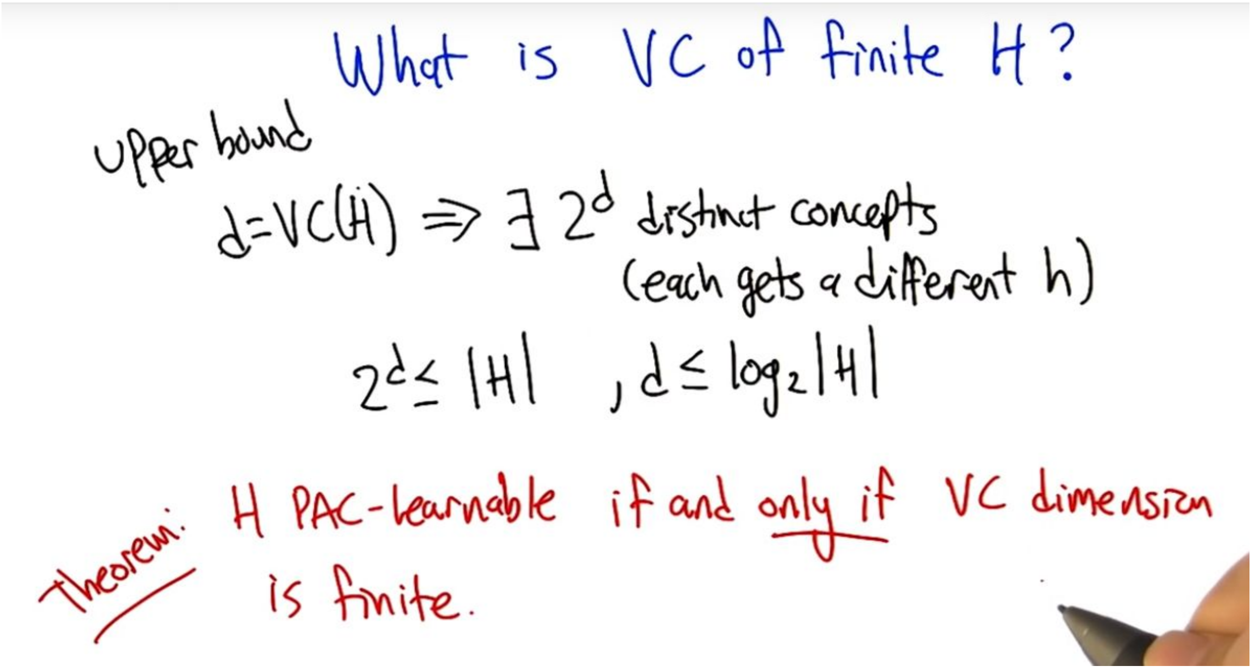
VC = $\boxed{}$



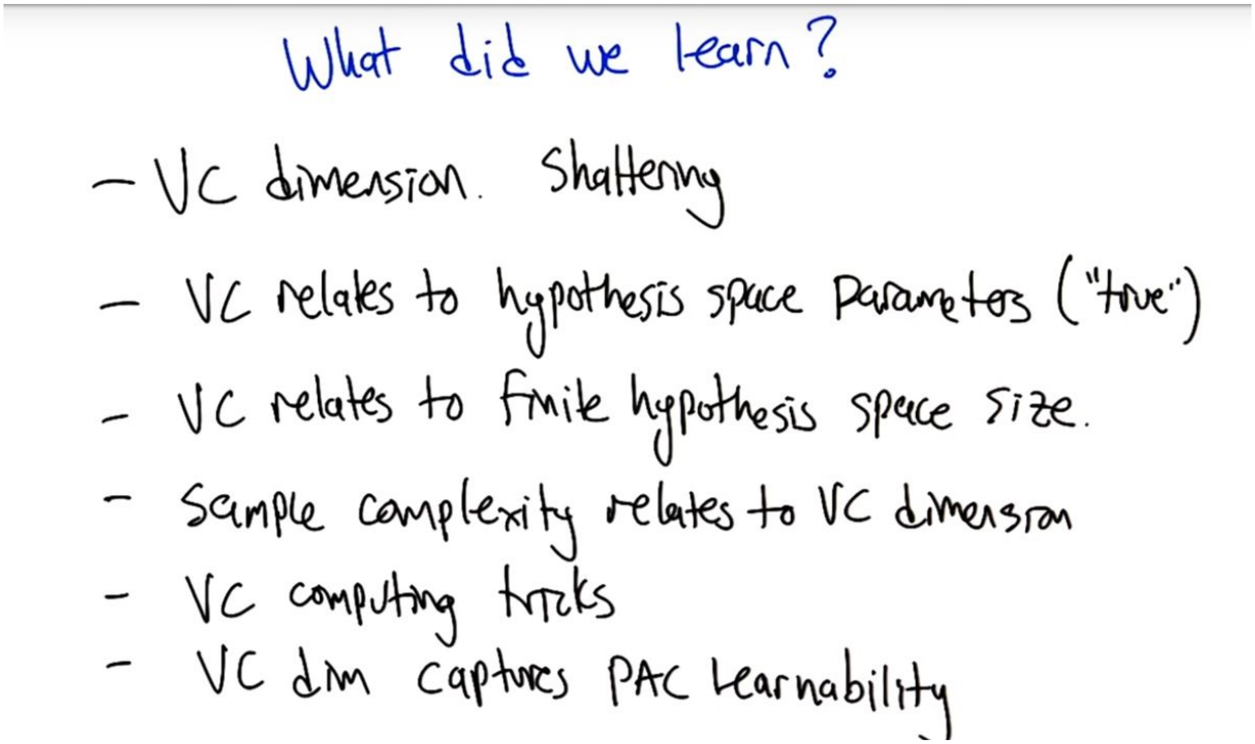
10. Sample Complexity



11. VC of Finite H



12. Summary


What did we learn?
- VC dimension. Shattering
- VC relates to hypothesis space parameters ("true")
- VC relates to finite hypothesis space size.
- Sample complexity relates to VC dimension
- VC computing tricks
- VC dim captures PAC learnability