Advanced Machine Learning: Problem Set II Solutions

Problem 1: Let $\mathcal{H}_{rec}^d$ be the class of axis aligned rectangles in $\mathbb{R}^d$. We have seen that VCdim($\mathcal{H}_{rec}^d$) = 4. Prove that in general $\mathcal{H}_{rec}^d = 2d$.

Proof:

(1) Show that the VC–dim $\geq 2d$.

Consider a set of 2d points where each point only has one of the d dimensions set to either 1 or -1 and 0 for all other dimensions. It is easy to see that any subset of these points can be shattered by an axis-aligned rectangle. Hence the VC–dim is atleast $2d$.

(2) Show that the VC–dim $< 2d + 1$.

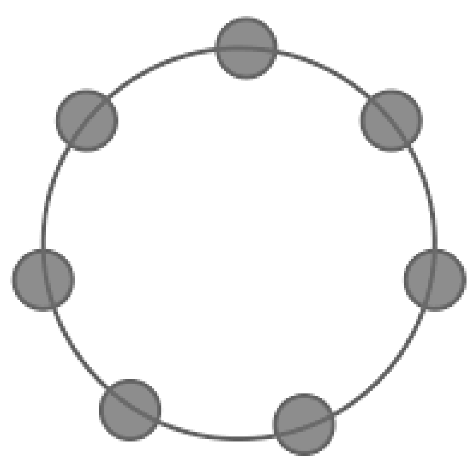
Consider a set of $2d + 1$ points. Consider finding the minimum and maximum of value in each dimension for these set of points and then building a R d rectangle with these bounds. Since there are 2d + 1 points, atleast one point must lie inside this rectangle. If we label this interior point as negative then there is no rectangle that can separate this labeling. This proves that VC–dim $< 2d + 1$.

Combining (1) and (2) we get that the VC–dim $= 2d$.

Problem 2: Let $\mathcal{H}_{tri}$ be the class of triangles in $\mathbb{R}^2$. Prove that $\mathcal{H}_{tri} = 7$.

Proof:

The VC-dimension of a triangle is at least 7. All possible labelling of the seven points aligned on a circle can be separated using the triangles. See the figure below.



Given 7 points on a circle, they can be labeled in any desired way because in any labeling, the negative examples form at most 3 contiguous blocks. Therefore one edge of the triangle can be used to cut off each block. However, no set of 8 points can be shattered. If one of the points is inside the convex hull of the rest, then it is not possible to label that point negative and the rest positive. Otherwise, it is also not possible to label them in alternating $+$, $-$, $+$, $-$, $+$, $-$, $+$, $-$ order.