Foundations of Data Science & Machine Learning

Tutotial 05

March 26, 2021

Question 1. An axis aligned rectangle in $\mathbb{R}^2$ is a rectangle whose sides are parallel to the x or y axes of the plane. The hypothesis corresponding to an axis aligned rectangle R is the function $f_R : X \to \{+1, -1\}$ which assigns $+1$ to all points inside R (including the boundary) and -1 to the points outside of R. Let H be the hypothesis class defined as

$$H = \{f_R : \ R \text{ is an axis-aligned rectangle in } \mathbb{R}^2\}.$$

Show that the VC dimension of H is 4.

Q1. In order to show that the VC dimension of H is 4, we will first show that it is at least 4 and then show that it is less than 5.

To show that it is at least 4 we are allowed to pick some special set S of 4 points and show that it can be shattered by H. Showing that S can be shattered involves showing any of the 16 possible labellings of S, there exists an axis aligned rectangle which contains all the points labelled $+1$ and no point labelled -1. With some trial and error, one can show that the set of 4 points $\{(0, 1), (0, -1), (1, 0), (-1, 0)\}$ can be be shattered by axis aligned rectangles. Hence VC dimension of H is at least 4.

To show that it is less than 5, we have to show that H does not shatter *any* set of 5 points and hence the set of S of 5 points has to be kept arbitrary. But on our side, we can pick any of the 32 possible labellings which will beat H. Consider the bounding box R of S, that is, the smallest axis aligned rectangle which contains all points in S. If any point of S is in the interior of R, then label it with -1. Label the points of S on the boundary of R as $+1$. One can see that any axis-aligned rectangle which contains all the $+1$'s will have to include the -1's also and hence this labelling cannot be realised by H. If no point is on the interior of R, then even after identifying one point in each side of R, we will get at least one point $s \in S$ free. Labeling s with -1 and others with $+1$ gives us the unshatterable configuration.

Question 2. Show that

1. Let $S = \{x_1, \ldots, x_n\}$ be a collection of n vectors in $\mathbb{R}^d$ which are linearly dependent. Moreover, the last coordinate is 1 for all the vectors. Show that S can be partitioned into two sets G and B such that the convex hulls of G and B have a non-empty intersection.

2. Show that any set S of $d+2$ points in $\mathbb{R}^d$ can be partitioned into two sets G and B such that the convex hulls of G and B have a non-empty intersection.

3. Show that the VC dimension of linear separators in $\mathbb{R}^d$ is at most $d+1$.

Q2.1. Since S is linearly independent, there exists coefficients $\lambda_1, \ldots, \lambda_n$ (not all zero) such that $\sum_{i=1}^{n} \lambda_i x_i = 0$. By taking all the terms with a negative coefficient to the RHS, we get

$$\sum_{x_i \in G} \lambda_i x_i = \sum_{x_i \in B} \lambda_i' x_i, \tag{1}$$

where $\lambda_i' = -\lambda_i$. Notice that G and B form a parition of S and the linear combinations on both sides of Equation 1 are non-negative linear combinations (no coefficient is negative). Since the last coordinate in each x_i is 1, reading the last cooridinate in Equation 1 gives $\sum_{x_i \in G} \lambda_i = \sum_{x_i \in B} \lambda_i'$ (call this number μ). Dividing all the coefficients in Equation 1 with this number makes both sides of Equation 1 a convex combination

$$\sum_{x_i \in G} \mu_i x_i = \sum_{x_i \in B} \mu_i' x_i = y, \tag{2}$$

where $\mu_i = \lambda_i/\mu$ and $\mu_i' = \lambda_i'/\mu$. Hence y is a point in the convex hull of both G and B.

Q2.2. Let S' be the set of points in $\mathbb{R}^{d+1}$ obtained by appending a 1 to every point in S. Since any set of $d+2$ points over $\mathbb{R}^{d+1}$ is linearly dependent, we get the required result from the previous question. We can ignore the padded 1 once the partition into G and B are obtained.

Q2.3. It follows from the previous question that no $(d+2)$-sized set in $\mathbb{R}^d$ can be shattered by hyperplanes. The labelling which gives $+1$ to all the points in G and -1 to all the points in B obtained as above cannot be separated by a hyperplane since their convex hulls overlap.