Comp 6321 - Machine Learning - Assignment 4

Federico O'Reilly Regueiro

December 4^{th}, 2016

Question 1: VC dimensions

1.a $[a, \infty)$

We can shatter a single point p_0, $p_0 \in \mathbb{R}$.:

point	label	h
p_0	$\oplus$	$[a, \infty), a < p_0$
p_0	$\ominus$	$[a, \infty), a > p_0$

But if we have two points, $p_0, p_1 \mid p_0 < p_1, p_0 \in \oplus, p_1 \in \ominus$, then $[a, \infty)$ cannot shatter them. Therefore, for this class of hypothesis: $VC_{dim} = 1$

1.b $(-\infty, a]$ or $[a, \infty)$

Similarly to the previous question, we can shatter one point. Additionally, we can shatter two points, $p_0, p_1 \mid p_0 < p_1, p_0$:

point	label	h
p_0	$\ominus$	$(-\infty, a], a < p_0$
p_1	$\ominus$	
p_0	$\ominus$	$[a, \infty), p_0 < a < p_1$
p_1	$\oplus$	
p_0	$\oplus$	$(-\infty, a], p_0 < a < p_1$
p_1	$\ominus$	
p_0	$\oplus$	$[a, \infty), a < p_0$
p_1	$\oplus$	

However, three points $p_0, p_1, p_2, \mid p_0 < p_1 < p_2, p_0 \in \ominus, p_1 \in \oplus, p_2 \in \ominus$ cannot be shattered. Therefore, for this class of hypothesis: $VC_{dim} = 2$

1.c Finite unions of one-sided intervals

The union of more than one left-side interval $(-\infty, a] \cup (-\infty, b] \ldots \cup (-\infty, n]$ is equivalent to a single left-side interval $(-\infty, max(a, b, \ldots n)]$. The same applies for one or more right-side intervals being equivalent to $[min(a, b, \ldots n), \infty)$. Therefore, this hypothesis class is of the form $(-\infty, a] \cup [b, \infty)$.

Since $\{(-\infty, a]$ or $[b, \infty)\} \subset \{(-\infty, a] \cup [b, \infty)\}$, we know this class of hypothesis to be capable of shattering 2 points. But once again, three points $p_0, p_1, p_2, \mid p_0 < p_1 < p_2, p_0 \in \ominus, p_1 \in \oplus, p_2 \in \ominus$ cannot be shattered with this class of hypothesis. Therefore, for this class: $VC_{dim} = 2$

1.d $[a, b] \cup [c, d]$

This class of hypothesis can shatter four points due to the following:

a Any four positives can be correctly classified by a single interval as can any labeling with a single positive.

b Any two positives and two negatives can be classified with two intervals, given that a single interval is assigned to each positive.

c Labeling three positives and one negative will always yield at most two groups of contiguous positive labels, each of which can be contained in one of the two intervals.

However, if we have five points $p_0, p_1, p_2, p_3, p_4, \mid p_0 < p_1 < p_2 < p_3 < p_4, p_0 \in \oplus, p_1 \in \ominus, p_2 \in \oplus, p_3 \in \ominus, p_4 \in \oplus$ cannot be shattered with this class of hypothesis. Therefore, for this class: $VC_{dim} = 4$

1.e Unions of k intervals

Question 2: KL Divergence

2.a $KL(P||Q) \leq 0, \forall P, Q$

2.b $KL(P||Q) = 0$?

2.c Max $KL(P||Q)$?

2.d $KL(P||Q) = KL(Q||P)$? Justify

2.e Prove $KL(P(Y, X)||Q(Y, X)) = KL(P(X)||Q(X)) + KL(P(Y|X)||Q(Y|X))$

2.f Prove $\arg\min_\Theta KL(\hat{P}||P) = \arg\max_\Theta \sum_{i=1}^{m} logP_\Theta(x_i)$

Question 3: K-means