Bogdan Alexe
Machine Learning
Faculty of Mathematics and Computer Science
University of Bucharest

Assignment 1

Due: Sunday, 21 April 2019

Problem 1 (0.5 points) - problem 2.5 in the book "Foundation of the Machine Learning", 2nd Edition

Let $\mathcal{X} = \mathbf{R}^2$ and consider the set of concepts defined by the area inside a right triangle ABC with the two catheti AB and AC parallel to the axes and with AB/AC = α (fixed constant > 0). Show, using similar methods to those used in the seminar class for the axis-aligned rectangles, that this class can be PAC-learned from training data of size $m \geq (3/\varepsilon)\ln(3/\delta)$.

Problem 2 (0.5 points) - problem from lecture 5.

Consider $\mathcal{H}_{\text{balls}}$ to be the set of all balls in $\mathbf{R}^2$: $\mathcal{H}_{\text{balls}} = \{B(x,r),\ x \in \mathbf{R}^2,\ r \geq 0\}$, where $B(x,r) = \{y \in \mathbf{R}^2 \mid \| y - x \|_2 \leq r\}$

What are the conditions for which a set A in $\mathbf{R}^2$ of size 4 is shattered by $\mathcal{H}_{\text{balls}}$? Justify your answer.

Problem 3 (0.75 points) – problem 5.2 in the book "Understanding Machine Learning: From Theory to Algorithms"

Assume you are asked to design a learning algorithm to predict whether patients are going to suffer a heart attack. Relevant patient features the algorithm may have access to include blood pressure (BP), body-mass index (BMI), age (A), level of physical activity (P), and income (I). You have to choose between two algorithms; the first picks an axis aligned rectangle in the two dimensional space spanned by the features BP and BMI and the other picks an axis aligned rectangle in the five dimensional space spanned by all the preceding features.

1. Explain the pros and cons of each choice.
2. Explain how the number of available labeled training samples will affect your choice

Problem 4 (0.75 points) - problem 6.3 in the book "Understanding Machine Learning: From Theory to Algorithms"

Let $\mathcal{X}$ be the Boolean hypercube $\{0,1\}^n$. For a set $I \subseteq \{1, 2, \ldots, n\}$ we define a parity function h_I as follows. On a binary vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$,

$$h_I(x) = \left(\sum_{i \in I} x_i\right) \bmod 2$$

(That is, h_I computes parity of bits in I .) What is the VC-dimension of the class of all such parity functions, $\mathcal{H}_{\text{n-parity}} = \{h_I : I \subseteq \{1, 2, \ldots, n\}\}$? Prove your claim.

Problem 5 (1 point) - problem 6.2 in the book "Understanding Machine Learning: From Theory to Algorithms"

Given some finite domain set, $\mathcal{X}$, and a number $k \leq |\mathcal{X}|$, figure out the VC-dimension of each of the following classes (and prove your claims):

1. $\mathcal{H}^{\mathcal{X}}_{=k} = \{h \in \{0,1\}^{\mathcal{X}} : |\{x : h(x)= 1\}| = k\}$: that is, the set of all functions that assign the value 1 to exactly k elements of $\mathcal{X}$.
2. $H_{\text{at-most-k}} = \{h \in \{0,1\}^{\mathcal{X}} : |\{x : h(x)= 1\}| \leq k \text{ or } |\{x : h(x) = 0\}| \leq k\}$.

Bonus Problem (1 point)

Compute the VC-dimension of the class of convex d-gons (convex polygons with exactly d sides) in the plane. Provide a detailed proof of your result.