Introduction to Machine Learning Course

Short HW3 – SVM, Optimization, and PAC learning

Submitted individually by Sunday, 11.06, at 23:59.

You may answer in Hebrew or English and write on a computer or by hand (but be clear).

Please submit a PDF file named like your ID number, e.g., 123456789.pdf.

Bonus (maximal grade is 100): Writing on a computer (using LyX/LaTeX, Word + Equation tool, etc.) = 2 pts.

1. **VC-dimension exercises:**

   Define the hypothesis class of axis aligned squares in $\mathbb{R}^2$.

   $$\mathcal{X} = \mathbb{R}^2,\ \ \mathcal{H}_{\text{sqr}} = \{h_\theta \mid \theta = (a, b, r) \in \mathbb{R}^3 \text{ s.t. } r > 0\}$$

   where a single hypothesis is defined by $h_\theta(\boldsymbol{x}) = \begin{cases} +1, & (a \le x_1 \le a + r) \wedge (b \le x_2 \le b + r) \\ -1, & \text{otherwise} \end{cases}$.

   Find $\text{VCdim}(\mathcal{H}_{\text{sqr}})$ and prove your answer (rigorously).

2. Let $K_1(u, v) = \langle \phi_1(u), \phi_1(v) \rangle,\ K_2(u, v) = \langle \phi_2(u), \phi_2(v) \rangle$ be two kernels with corresponding feature mappings $\phi_1 : \mathcal{X} \to \mathbb{R}^{n_1}, \phi_2 : \mathcal{X} \to \mathbb{R}^{n_2}$ where $n_1, n_2 \in \mathbb{N}$. Notice that K is a valid (i.e., well-defined) kernel since it can be written as an inner product of a mapping of u and v.

   Prove that $K_3(u, v) = 2 \cdot K_1(u, v) + 3 \cdot K_2(u, v)$ is a valid kernel. That is, propose a feature mapping $\phi_3 : \mathcal{X} \to \mathbb{R}^{n_3}$ for some $n_3 \in \mathbb{N}$, such that $K_3(u, v) = \langle \phi_3(u), \phi_3(v) \rangle$.

3. We will now prove that the following Soft-SVM problem is convex:

   $$\boxed{\underset{w \in \mathbb{R}^d}{\operatorname{argmin}} \frac{1}{m} \sum_{i=1}^{m} \max\{0, 1 - y_i \cdot w^\top x_i\} + \lambda \|w\|_2^2}$$

   Let $f, g : C \to \mathbb{R}$ be two convex functions defined over a convex set C.

   Lemma (no need to prove)**:** given a constant $\alpha \in \mathbb{R}_{\ge 0}$, the function $\alpha f(z)$ is convex w.r.t z.

   Lemma (no need to prove)**:** a sum of any number of convex functions is convex.

   3.1. Prove (by definition) that $q(z) \triangleq \max\{f(z),\ g(z)\}$ is convex w.r.t z.

   3.2. Using a rule from Tutorial 07, conclude that $\max\{0, 1 - y_i \boldsymbol{w}^\top \boldsymbol{x}_i\}$ is convex w.r.t $\boldsymbol{w}$.

   3.3. Using the above (and properties from Tutorial 07), conclude that the Soft-SVM optimization problem is convex w.r.t $\boldsymbol{w}$.