

1 Εκμάθηση Ορθογωνίων [2 μονάδες]

Η κλάση H των ορθογωνίων διαχωριστών σε d διαστάσεις περιέχει συναρτήσεις της μορφής:

$$h_{a,b}(x) = \begin{cases} +1 & \text{Αν για κάθε } i \in [d], \text{ ισχύει } x_i \in [a_i, b_i] \\ -1 & \text{διαφορετικά} \end{cases}$$

για παραμέτρους $a_1, \dots, a_d, b_1, \dots, b_d \in \mathbb{R}$. Δηλαδή, οι συναρτήσεις παίρνουν την τιμή $+1$ αν το σημείο είναι μέσα στο ορθογώνιο και -1 αλλιώς.

- (α) Να δείξετε ότι στη realizable περίπτωση, ο αλγόριθμος A που για ένα σύνολο δειγμάτων βρίσκει το μικρότερο ορθογώνιο που περιέχει τα θετικά σημεία είναι ERM.
- (β) Να δείξετε ότι αν ο A είναι (ϵ, δ) -realizable PAC learner με $m_H^{\text{PAC}} = O\left(\frac{d}{\epsilon} \log(d/\delta)\right)$.
Συμβουλή: Δείξτε ότι ο αλγόριθμος A επιστρέφει ένα ορθογώνιο με παραμέτρους \hat{a}, \hat{b} που είναι υποσύνολο του πραγματικού ορθογωνίου h_{a^*, b^*} . Στη συνέχεια δείξτε ότι για κάθε $i \in [d]$, $Pr[x_i \in (a_i^*, \hat{a}_i)] < \epsilon/(2d)$ και $Pr[x_i \in (\hat{b}_i, b_i^*)] < \epsilon/(2d)$ με μεγάλη πιθανότητα.
- (γ) Υπολογίστε το VC dimension της κλάσης H .
- (δ) Ποια είναι τα bounds της δειγματικής πολυπλοκότητας για realizable και agnostic PAC learning που συνεπάγονται για τον αλγόριθμο A ;

2 Εκμάθηση από 2 πηγές δειγμάτων [1 μονάδα]

Εξετάστε μια παραλλαγή του μοντέλου PAC στην οποία υπάρχουν δύο πηγές δειγμάτων: η μια δίνει θετικά παραδείγματα και η άλλη αρνητικά, και τα δύο σύμφωνα με μια υποκείμενη κατανομή D στο X .

Πιο συγκεκριμένα, για μια συνάρτηση $f : X \rightarrow \{-1, +1\}$, έστω D_+ η κατανομή πάνω στο $X_+ = \{x \in X : f(x) = +1\}$ με μάζα πιθανότητας $D_+(A) = D(A)/D(X_+)$, για κάθε $A \subseteq X_+$. Ομοίως, D_- είναι η κατανομή στο $X_- = \{x \in X : f(x) = -1\}$. Ο ορισμός του PAC learnability στο μοντέλο δύο πηγών είναι ο ίδιος με τον standard ορισμό του PAC learnability εκτός από το ότι εδώ ο εκπαιδευόμενος έχει πρόσβαση σε $m_+(\epsilon, \delta)$ i.i.d. παραδείγματα από τα D_+ και $m_-(\epsilon, \delta)$ i.i.d. παραδείγματα από το D_- . Ο στόχος είναι να προσδιοριστεί μια h τ.ω. με πιθανότητα τουλάχιστον $1 - \delta$, $Pr_{x \sim D_+}[h(x) \neq +1] \leq \epsilon$ και $Pr_{x \sim D_-}[h(x) \neq -1] \leq \epsilon$.

- (α) Δείξτε ότι εάν μια κλάση υποθέσεων H είναι PAC learnable στο τυπικό μοντέλο μιας πηγής παραδειγμάτων, τότε είναι PAC learnable και στο μοντέλο δύο πηγών.
- (β) Έστω h_+ και h_- οι συναρτήσεις που είναι πάντα $+1$ και πάντα -1 αντίστοιχα. Υποθέτοντας ότι $h_+, h_- \in H$, δείξτε ότι εάν η κλάση H είναι PAC learnable στο μοντέλο δύο πηγών παραδειγμάτων, τότε είναι PAC learnable και στο τυπικό μοντέλο μιας πηγής.