Домашнее задание №3 по курсу «Машинное обучение»: основы машинного обучения

Скавыш Максим

Задание 1

Вычислите VCdim(H), если H — семейство линейных бинарных классификаторов в d-мерном пространстве.

Решение

Зафиксируем некоторое n – произвольная размерность пространства. Будем рассматривать семейство линейных бинарных классификаторов вида (bias можно внести в w):
$H = sign \circ h_w, h_w = sign(\langle w, x\rangle)$, где $w \in R^n,\ x \in R^n, b \in R$
Докажем существование множества $C \in X$ размерности n, которое можно раскрасить. Пусть $C = \{c_i = e_i, i = 1, \dots ,n\}$, где $e_i = (0, \dots ,\ 0, 1, 0, \dots , 0)$ – единичные вектора с i-м ненулевым элементом. Такое множество C можно раскрасить подбирая w так, что $\langle w, e_i\rangle = y_i$. Тоесть существует множество размерности n которое можно раскрасить.
Докажем что невозможно раскрасить любое n + 1 множество
Пусть $C = \{c_i,\ i = 1, \dots , n + 1\}$. Так как размерность пространства n: $c_i \in R^n$, то как минимум один элемент из множества C линейно выражается через остальные элементы. Для определенности считаем что это $c_{n+1} = \sum_{i=1}^{n} \alpha_i c_i$. Тогда и y_i однозначно определяется из линейной зависимости ⇒ мы не можем раскрасить произвольное множество C из n+1 элемента.
$\mathrm{VCdim}(\mathrm{H_w}) = \mathrm{n}$

Задание 3

Пусть X — булев гиперкуб размерности n. Для множества $I \in \{1, 2, \dots , n\}$ и объекта $x \in X$, $x = (x_1, x_2, \dots , x_n)$ зададим функцию $h_I(x) = (\sum\limits_{i \in I} x_i) \mod 2$. Чему равна VCdim таких множества всех таких функций?

Решение

Зафиксируем некоторое произвольное n. Тогда размер множества $|X| = 2^n$. Известно:
$VCdim(H) \le \log_2|H| = n$
Докажем что существует множество из n элементов которое можно разукрасить:
$C = \{x^i = \left(x_1^i, \dots , x_n^i\right), i = 1, \dots , n$
Возьмем элемент $x^i \in C\ i \in [1, n]$ таким что все его элементы кроме i-го нулевые:
$x^i = \left(x_1^i, \dots , x_i^i \dots , x_n^i\right) = (0, \dots , 1 , \dots , 0)$ тогда добавляя в множество I те i для которых мы хотим получить $y_i = 1$ а для остальных беря $I = \emptyset$. Мы всегда можем разукрасить множество размерности n. Получили $\mathrm{VCdim}(\mathrm{H}_I) = \mathrm{n}$

Задание 4

Объясните, как согласуются:

• ERM-алгоритм над конечным классом H — PAC-learnable в случае гипотезы реализуемости и No Free Lunch theorem?
• ERM-алгоритм над конечным классом H — agnostic PAC-learnable и No Free Lunch theorem?

Решение

При NFL теореме мы заранее фиксируем некоторое значение $m \leq \frac{|X|}{2}$ и говорим что при выборке размера m существует такое распределние D что выполяnются условия .