

1. Напишите определения, что означают слова «*PAC*-обучаемый» и «эффективно *PAC*-обучаемый»?
2. Рассмотрим задачу распознавания чёрно-белого текста (например, на русском языке). Имеются образцы изображений всех букв. На вход подаются отдельные буквы, заданные как двумерные массивы размера $m \times n$, заполненные нулями и единицами. Что будет являться в данной задаче признаковым пространством, представлением, классом представлений?
3. В задаче поиска прямоугольника мы получили такую оценку: $4(1 - \frac{\varepsilon}{4})^m \leq \delta$. Докажите, что m полиномиально зависит от ε и δ , то есть существует такой полином $p(x, y)$, что $m(\varepsilon, \delta) \leq p(\frac{1}{\varepsilon}, \frac{1}{\delta}) \forall \varepsilon, \delta : 0 < \varepsilon, \delta < 1$. Используйте неравенство $1 - x \leq e^{-x}$, верное для всех $x > 0$.
4. Рассмотрим одномерный вариант задачи с прямоугольником: пусть задан некоторый отрезок c , распределение D на \mathbb{R} и оракул $EX(c, D)$, выбирающий произвольную точку p из D и возвращающий $(p, 1)$, если $p \in c$, и $(p, 0)$ иначе. Алгоритм делает m вызовов оракула и возвращает отрезок c' . Пусть $0 < \varepsilon, \delta < 1$. Каким должен быть c' и как должно зависеть m от ε и δ , чтобы выполнялось условие *PAC*-обучаемости? Доказывать оценку не надо, только найти по аналогии с задачей о прямоугольнике.
5. Допустим, вы хотите проверить ваш любимый алгоритм машинного обучения на *PAC*-обучаемость. Напишите, какую задачу он решает, какими будут признаковое пространство, представления и класс представлений.