Домашнее задание №1 по курсу "Машинное обучение"

Прокопенко Тимофей

Задача 1.

Решение:
Из обратимости функции $f(t) = \delta$ следует существование $f^{-1}(\delta) = t$, где $\delta \in \mathbb{R}, t > 0$.
Учитывая, что для $\forall t > 0$ $\mathbb{P}[X > t] \leq f(t)$ и $\mathbb{P}[X > t]$ принимает значения от 0 до 1, можно ограничить область значения функции f до (0;∞), т.е. $\delta \in (0; \infty)$.

Проведем ряд преобразований:
$\mathbb{P}[X > t] \leq f(t) \iff 1 - \mathbb{P}[X \leq t] \leq f(t) \iff \mathbb{P}[X \leq t] \geq 1 - f(t)$.
Подставим $f(t) = \delta$ и $f^{-1}(\delta) = t$ в последнее выражение : $\mathbb{P}[X \leq f^{-1}(\delta)] \geq 1 - \delta, \forall \delta > 0$

Задача 2.

Решение:
а) Для начала рассмотрим классификатор h_s на такой выборке из S (назовем ее S_1), где все $y_i = 1$.
То есть

$$h_S(x) = \begin{cases} 1, & \text{если } x \in S_1 \\ 0, & \text{иначе} \end{cases}$$

Для того, чтобы доказать, что в h_p найдется классификатор, совпадающий с h_s, нужно найти такой полином p(x), который принимает отрицательные значения, если $x \notin S_1$, а при $x \in S_1$ его значения неотрицательны.

Пример такого полинома: p(x)=$-\prod\limits_{x_k \in S_1} (x - x_k)^2$. При всех $x \in \mathbb{R}$ p(x) принимает отрицательные значения, кроме точек из S_1, где полином обращается в ноль. Таким образом мы доказали, что в классе h_p существует классификатор, совпадающий с h_s.

б) Мы знаем, что ERM-гипотеза h_s приводит к переобучению. В пункте а) было доказано, что в h_p найдется классификатор, совпадающий с h_s, таким образом применение ERM-парадигмы для h_p может привести к переобучению.

Задача 3.

Решение:
Из предположения о реализуемости можно сделать вывод, что существует такая гипотеза h^*, алгоритм которой выбирает прямоугольник, содержащий все точки положительного класса и только их.

Алгоритм A выбирает наименьший прямоугольник, содержащий все точки положительного класса. Очевидно, что этот прямоугольник лежит внутри прямоугольника, выбираемого алгоритмом гипотезы h^*. Таким образом, внутри прямоугольника, выбираемого алгоритмом A, не будет точек отрицательного класса, значит эмпирический риск будет равен нолю, т.е. A является реализацией ERM-алгоритма.

Задача 4.

Решение:
Нам необходимо выбрать радиус r таким образом, что ERM(h_r) = min(ERM(h)). Алгоритмы поиска данной ERM-гипотезы будет заключаться в следующем:

1. Рассчитаем все радиусы окружностей, проходящих через точки обучающей выборки ($r_i = x_{i1}^2 + x_{i2}^2$).
2. Выбираем радиус, порождающий классификатор, на котором ERM-ошибка минимальна.

Данный алгоритм будет являться полиномиальным, так как ERM-ошибка зависит от гипотезы h_r, гипотеза h_r зависит от выбранного радиуса r, а соответствующий радиус мы выбираем исходя из объектов обучающей выборки.