
Machine Learning, 2023 Spring

Assignment 1

Notice

Plagiarizer will get 0 points.

\LaTeX is highly recommended. Otherwise you should write as legibly as possible.

Problem 1 If $\mu = 0.9$, what is the probability that a sample of 10 marbles will have $\nu \leq 0.1$?

[Hints: 1. Use binomial distribution. 2. The answer is a very small number]

Problem 2 If $\mu = 0.9$, use the Hoeffding Inequality to bound the probability that a sample of 10 marbles will have $\nu \leq 0.1$ and compare the answer to the previous exercise.

Problem 3 We are given a data set \mathcal{D} and of 25 training examples from an unknown target function $f : \mathcal{X} \rightarrow \mathcal{Y}$, where $\mathcal{X} = \mathbb{R}$ and $\mathcal{Y} = \{-1, +1\}$. To learn f , we use a simple hypothesis set $\mathcal{H} = \{h_1, h_2\}$ and, where h_1 is the constant $+1$ function and h_2 is the constant -1 .

We consider two learning algorithms, S (smart) and C (crazy). S chooses the hypothesis that agrees the most with \mathcal{D} and C chooses the other hypothesis deliberately. Let us see how these algorithms perform out of sample from the deterministic and probabilistic points of view. Assume in the probabilistic view that there is a probability distribution on \mathcal{X} , and let $\mathbb{P}[f(x) = +1] = p$.

- (a) Can S produce a hypothesis that is guaranteed to perform better than random on any point outside \mathcal{D} ?
- (b) Assume for the rest of the exercise that all the examples in \mathcal{D} have $y_n = +1$. Is it possible that the C hypothesis that produces turns out to be better than the hypothesis that S produces?
- (c) If $p = 0.9$, what is the probability that S will produce a better hypothesis than C ?
- (d) Is there any value of p for which it is more likely than not that C will produce a better hypothesis than S ?

Problem 4 A friend comes to you with a learning problem. She says the target function f is completely unknown, but she has 4,000 data points. She is willing to pay you to solve her problem and produce for her a g which approximate f . What is the best that you can promise her among the following:

- (a) After learning you will provide her with a g that you will guarantee approximates f well out of sample.
- (b) After learning you will provide her with a g , and with high probability the g which you produce will approximate f well out of sample.
- (c) One of two things will happen.
 - (i) You will produce a hypothesis g ;
 - (ii) You will declare that you failed.

If you do return a hypothesis g , then with high probability the g which you produce will approximate f well out of sample.

Problem 5

Given target function $f(x) = ax$ (a unknown), now we have the dataset $\{x_1, y_1\} \dots \{x_n, y_n\}$. Which hypothesis class will you chose, $H = \{ax + b\}$ or $H = \{ax\}$, explain your reason.