ML.exer.chap11

samira moradzade

January 2022

1 xercise 11.1:

Consider a case in that the label is chosen at random according to $P[y = 1] = [y = 0] = \frac{1}{2}$ Consider a learning algorithm that outputs the constant predictor $h(x) = 1$ if the parity of the labels on the training set is 1 and otherwise the algorithm outputs the constant predictor $h(x) = 0$. Prove that the difference between the leave-one-out estimate and the true error in such a case is always $\frac{1}{2}$.

proof: first consider S set as a i.i.d sample
we know h as the out put of learning algorithm.
because h is a constant function we have $L_D(h) = \frac{1}{2}$ we want to calculate the $L_V(h)$. assume the parity of S is 1
then fix some fold $\{(x, y)\} \subseteq S$
we have two cases bellow:

FIRST:
as the $S \setminus \{X\}$ is 1 so $Y = 0$ and since trained using $S \setminus \{X\}$ the algorithm oytputs the predictor $h(x) = 1$
therefor the leave-one-out estimate using this fold is 1.

SECOND:
as the $S \setminus \{X\}$ is 0 so $Y = 1$ and since trained using $S \setminus \{X\}$ the algorithm oytputs the predictor $h(x) = 0$
therefor the leave-one-out estimate using this fold is 1.
after avvaraging the two folds we calculate we find out that the esstimate error of h is 1. and the different between estimation error and the true error is $\frac{1}{2}$.
for the other case (the parity be 0) analyze in the same way.