Lecture 9 - 07-04-2020

$\hat{h}isERMpredictor$

$$\ell_D\left(\hat{h}\right) \leq min\ \ell_D\left(h\right) + \sqrt{\frac{2}{m}\ \ln\frac{2\,H}{\delta}} \qquad with\ prob.\ at\ least\ 1-\delta$$

Now we do it with tree predictors

1.1 Tree predictors

$$X = \{0,1\}^d \longrightarrow \textbf{Binary classification}$$
$$h : \{0,1\}^d \longrightarrow \textbf{Binary classification H}1$$

How big is this class?
Take the size of codomain power the domain $\longrightarrow |H| = 2^{2^d}$
Can we have a tree predictor that predict every H in this class?
For every $h : \{0,1\}^d \longleftrightarrow \{-1,1\} \quad \exists T$

We can **build a tree** such that $\quad h_T = h$

Figure 1.1: Tree building

$X = (0, 0, 1, ..., 1) \qquad h(x) = -1$
$x_1, x_2, x_3, ..., x_d$

I can apply my analisys to this predictors
If I run ERM on H

$$\ell_D\left(\hat{h}\right) \leq min\, \ell_D\left(\hat{h}\right) + \sqrt{\frac{2}{m} 2^d \ln 2 + \ln \frac{2}{\delta}} \qquad \longrightarrow \ln|H| + \ln \frac{2}{\delta}$$

No sense! What we find about training set that we need?
Worst case of overfitting $m >> 2^D = |X| \Rightarrow$ training sample larger

PROBLEM: cannot learn from a class to big (H is too big)
I can control H just limiting the number of nodes.

$H_N \longrightarrow$ tree T with at most N node, $N << 2^D$
$|H_N| = ?$

$|H_N| = (\# \textit{ of trees with } \leq N \textit{ nodes}) \times (\# \textit{ of test on interval nodes }) \times (\# \textit{ labels on leaves})$

$$|H_N| = \otimes \times d^M \times 2^{N-M}$$

N of which $N - M$ are leaves

Figure 1.2: Tree with at most N node

⊗*# of binary trees with N nodes, called* ***Catalan Number***

1.1.1 Catalan Number

*We are using a binomial *

$$\frac{1}{N}\binom{2N-2}{N-1} \quad \leq$$