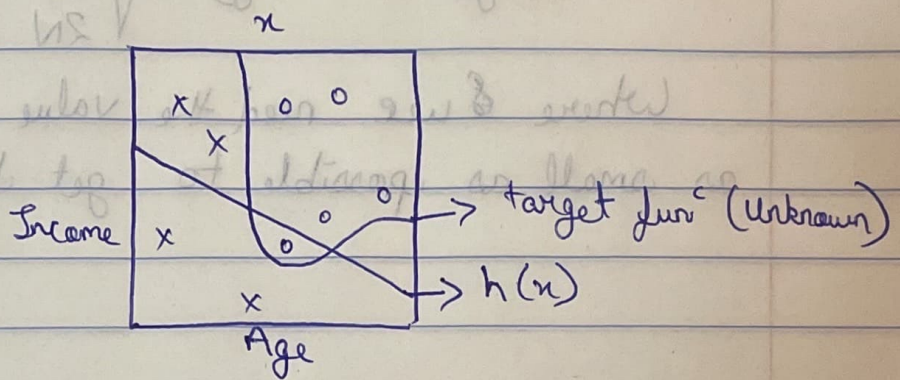
Lecture 3

Token no:- CNy2Tg

Report:-

The lecture started with initial recap to the learning problem setup, then moving ahead to give a short recap on Hoeffdings Inequality. We studied how the bin model is related to learning.



We try to find the $E_{in}(h)$ which is fraction of data points misclassified.

We calculate E_{in} with formula.

$$\nu = E_{in} = \frac{1}{N} \sum_{n=1}^{n=N} [h(x_n) \neq f(x_n)]$$

This will return 0 or 1 if condition is satisfied or not. i.e if the data are classified correctly or not.

New Hoeffdings Eq^n becomes;

$$P[|E_{in} - E_{out}| > \varepsilon] \leq 2e^{-2\varepsilon^2 N}$$

This condition is for a misclassification.

Our goal is to get as low as possible for E_{out} using our input sample E_{in}.
We then saw an example of Selection Bias Illustrated with Coin tossing.
Then we tried to find the Hoeffdings bound for a finite set of Hypothesis $\mathcal{H}$.

$$E_{out}(g) \leq E_{in}(g) + \sqrt{\frac{1}{2N} \log \frac{2|\mathcal{H}|}{\delta}}$$

Where we need the value of δ to be as small as possible to get high confidence.