1) PAC Learning (probably approximately correct)

δ - confidence par.

ε - accuracy par.

$m_H : (0,1)^2 \to N$

$L_{D,f}(h) \le \varepsilon$

2) $P_{x\sim D}[h^*(x) = f(x)] = 1$

D - probability distribution over $X \times Y$

$$L_D(h) = \underset{(x,y)\sim D}{P}[h(x) \ne y] = D(\{(x,y) : h(x) \ne y\})$$

2) Agnostic PAC

$$L_D(h) \le \min_{h \in H} L_D(h') + \varepsilon$$

3) Generalize loss function

$$L_D(h) = \underset{z\sim D}{E}[\ell(h,z)] \text{ - risk function}$$

$$L_S(h) = \frac{1}{m}\sum_{i=1}^{m} \ell(h, z_i)$$

4) Uniform convergence

ε-representative sample:

$$\forall h \in H \; |L_S(h) - L_D(h)| \le \varepsilon$$

If $\varepsilon/2$ - repr $\Rightarrow$ $h_S \in \underset{h\in H}{\operatorname{argmin}} L_S(h)$ satisfies

$$L_D(h_S) \le \min_{h\in H} L_D(h)$$

$$L_D(h_S) \le L_S(h_S) + \varepsilon/2 \le \underset{\wedge}{L_S}(h) + \varepsilon/2 \le L_D(h) + \varepsilon$$

m_H^{uc} : $L_D(h) + \varepsilon/2$

5) Finite classes are Agnostic PAC

$$D^m(\{S : \forall h \in H \; |L_S(h) - L_D(h)| \le \varepsilon\}) \ge 1 - \delta$$

$> \varepsilon$ $< \delta$

$$\bigcup_{h\in H} \{S : |L_S(h) - L_D(h)| > \varepsilon\} \le \sum_{h\in H} D^m(\{S : |L_S(h) - L_D(h)| > \varepsilon\}) \le$$

$$E(L_S(h)) = E\left[\frac{1}{m}\sum_{i=1}^{m} \ell(h, z_i)\right] = E[\ell(h,z)] = L_D$$

$$|L_S(h) - E(L_S(h))|$$

Use Hoefding's ineqv

$$\le \sum_{h\in H} 2\exp(-2m\varepsilon^2) = 2|H| \exp\left(-\frac{2m\varepsilon^2}{B^2}\right)$$

$$m \ge \frac{\log(2|H|/\delta)}{2\varepsilon^2}$$

6) Enjoys UC-property:

$$m_H^{uc}(\varepsilon, \delta) \le \frac{\log(2|H|/\delta)}{2\varepsilon^2}$$

H is pac-learnable

$$\Leftrightarrow \lim_{m\to\infty} \underset{S\sim D^m}{E}[L_D(A(S))] = 0$$