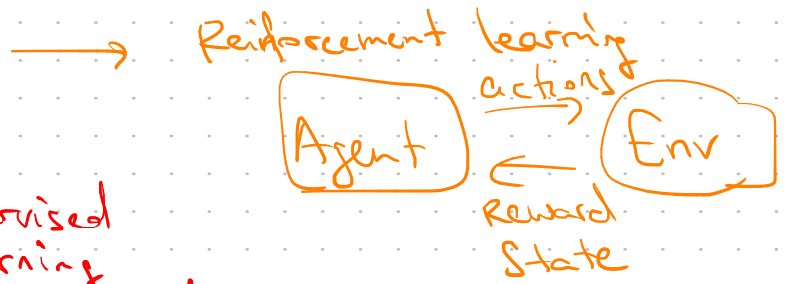


Machine Learning

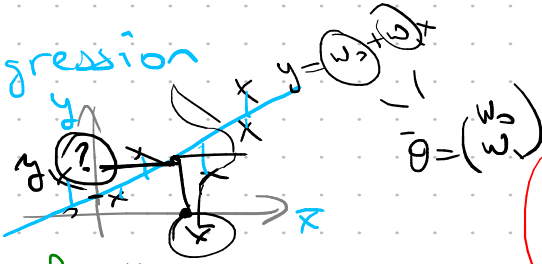


Supervised learning

$$D = \{(\bar{x}_n, \bar{y}_n)\}_{n=1}^N$$

$$\bar{x} \sim \text{[box]} \rightarrow \text{[circle]} \text{ } p(y|\bar{x})$$

Regression



Classification



Unsupervised learning

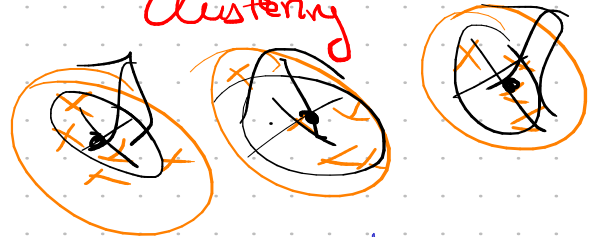
$$D = \{\bar{x}_n\}_{n=1}^N$$

$$p(\bar{x})$$

Representation learning



Clustering



$$D \rightsquigarrow \bar{\theta}$$

prior $p(\bar{\theta})$

likelihood $p(\bar{x}|\bar{\theta})$

likelihood

$$p(D|\bar{\theta}) = \prod_{n=1}^N p(\bar{x}_n|\bar{\theta})$$

posterior

$$p(\bar{\theta}|D) = \frac{p(\bar{\theta}) p(D|\bar{\theta})}{p(D)}$$

POSTERIOR

opēn / peluka
1 / 0

$$\theta = p(\text{opēn}|\theta)$$

$$D = \text{oppoopop}$$

$$p(D|\theta) = \theta^4 (1-\theta)^3 \rightarrow \max_{\theta}$$

$$\frac{\partial p(D|\theta)}{\partial \theta} = 4\theta^3(1-\theta)^3 - 3\theta^4(1-\theta)^2 = 0$$

$$1) \bar{\theta}_{ML} = \arg \max_{\bar{\theta}} p(D|\bar{\theta})$$

$$\bar{\theta}_{ML} = 4/7$$

maximum likelihood

$$= \theta^3(1-\theta)^2$$

$$4(1-\theta) - 3\theta = 0$$

$$D = \text{opēn} \quad p(D|\theta) = \theta \rightarrow \max_{\theta}$$

$$\bar{\theta}_{ML} = \infty$$

$$\theta = \frac{4}{7} \in [0, 1]$$

$$2) \quad p(\bar{\theta} | D) \propto p(\bar{\theta}) p(D | \bar{\theta}) = p(\bar{\theta}) \prod_{n=1}^N p(x_n | \bar{\theta})$$

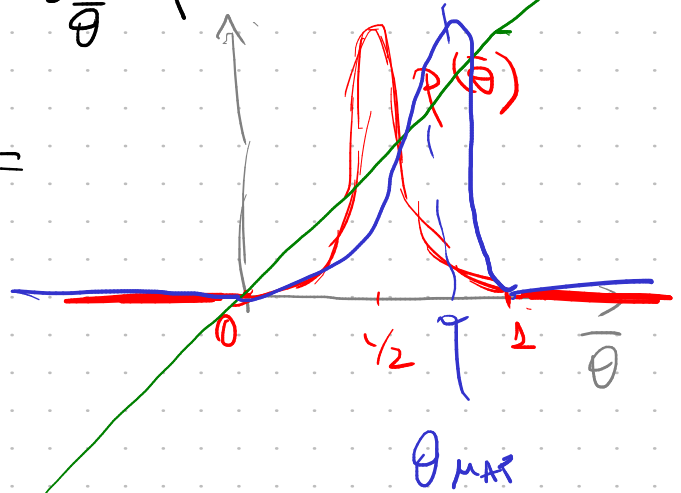
$$\bar{\theta}_{MAP} = \underset{\bar{\theta}}{\operatorname{argmax}} p(\bar{\theta} | D) = \underset{\bar{\theta}}{\operatorname{argmax}} p(\bar{\theta}) p(D | \bar{\theta})$$

maximum a posteriori

$$3) \quad p(\bar{x} | D) = \int p(x, \bar{\theta} | D) d\bar{\theta} =$$

$$= \int \underbrace{p(x | \bar{\theta})}_{\text{likelihood}} \underbrace{p(\bar{\theta} | D)}_{\text{posterior}} d\bar{\theta}$$

predictive distribution



$$p(\text{open} | D) = \int \overbrace{p(\text{open} | \theta)}^{\text{likelihood}} \overbrace{p(\theta | D)}^{\text{posterior}} d\theta$$