Machine learning

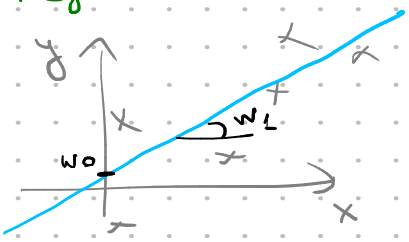
Supervised learning

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

$\bar{x}$ — y

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

↳ Regression



↳ Classification

Unsupervised learning

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

$p(\bar{x})$

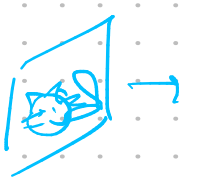
$\in \mathbb{R}^{3 \cdot 10^6}$

clustering

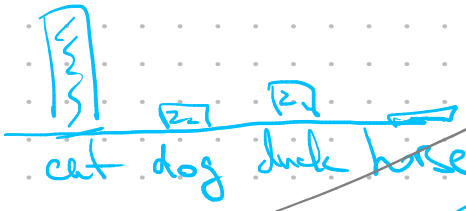
Reinforcement learning

Agent —act→ Env

rewards, states



→ cat / dog ? [0, 1]



$\theta = p(\text{орёл})$

$p(\text{решка}) = 1-\theta$

$D =$ ООРОРРРО $\rightsquigarrow \theta$

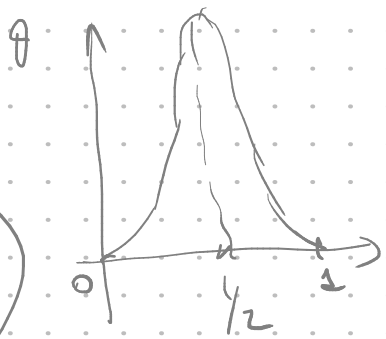
$p(y|x)p(x) = p(x,y) = p(x|y)p(y)$

$$p(y|x) = \frac{p(x|y)\,p(y)}{p(x)}$$

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

$p(\theta|D)$ — posterior; $p(D|\theta)$ — likelihood; $p(\theta)$ — prior; $p(D)$ — evidence

model params, data



$p(\theta|D) \propto p(D|\theta)\,p(\theta)$

$D = \{n \times \text{орёл}, m \times \text{решка}\} =$ ОО...Р...ОР

Likelihood:

$$p(D|\theta) = \theta^n (1-\theta)^m \xrightarrow[\theta]{} \max$$

$p(D|\theta=a) = a^n(1-a)^m$

maximum likelihood: $\theta_{ML} = \frac{n}{n+m}$

$$\frac{\partial p}{\partial \theta} = n\theta^{n-1}(1-\theta)^m - m\theta^n(1-\theta)^{m-1} =$$

$$= \theta^{n-1}(1-\theta)^{m-1}\left(n(1-\theta) - m\theta\right) = 0$$

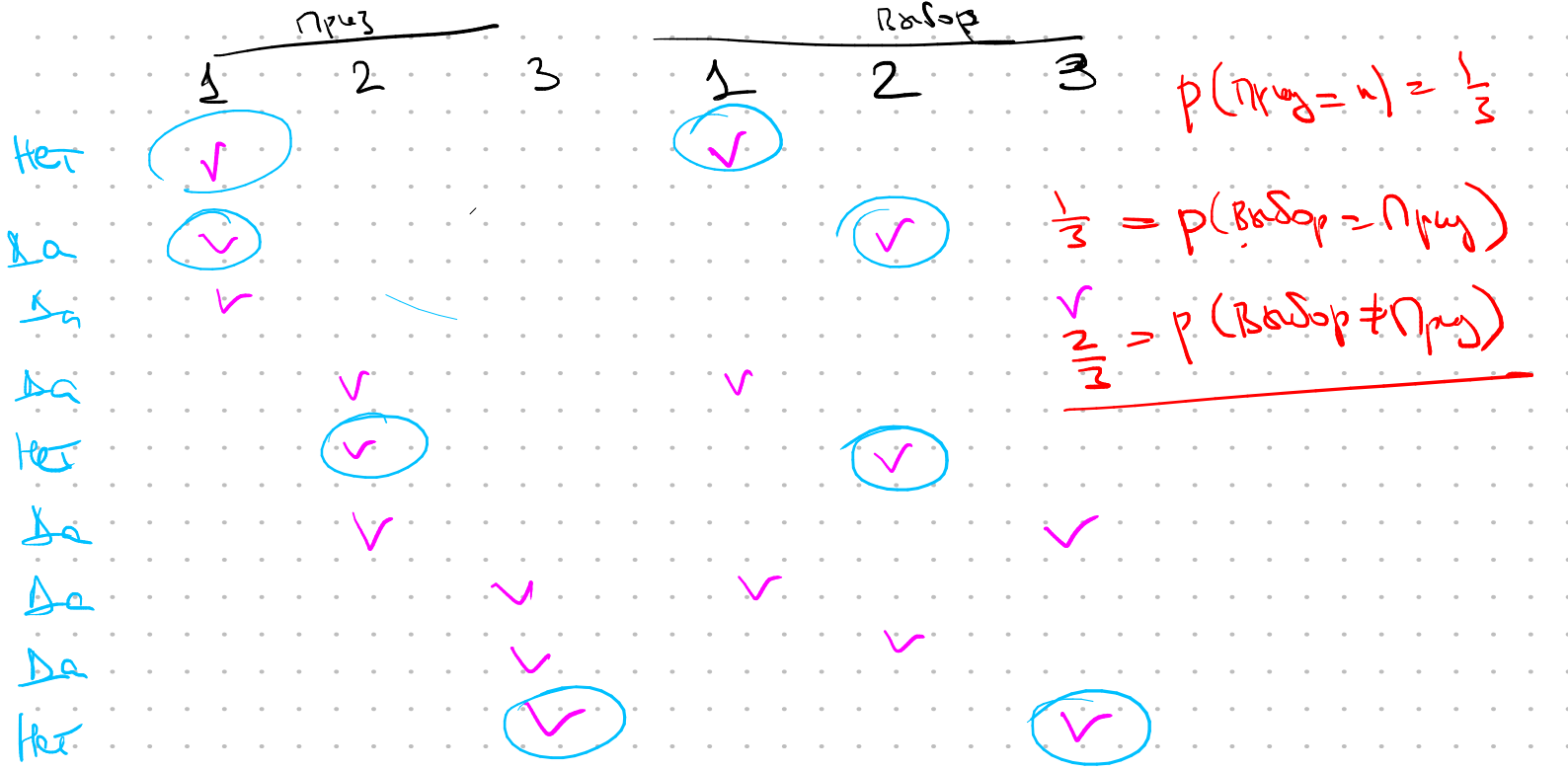
$\theta = 0,\ 1,\ \frac{n}{n+m}$

$D = $ ОО

$p(D|\theta) = \theta^2$

$\theta_{ML} \to \infty$ ~~$\theta_{ML} \to \infty$~~

$\theta_{ML} = 1$


Приз
Выбор
1
2
3
1
2
3
Нет
Да
Да
Нет
Да
Да
Да
Нет
$p(\text{Приз}=1) = \frac{1}{3}$
$\frac{1}{3} = p(\text{Выбор} = \text{Приз})$
$\frac{2}{3} = p(\text{Выбор} \neq \text{Приз})$