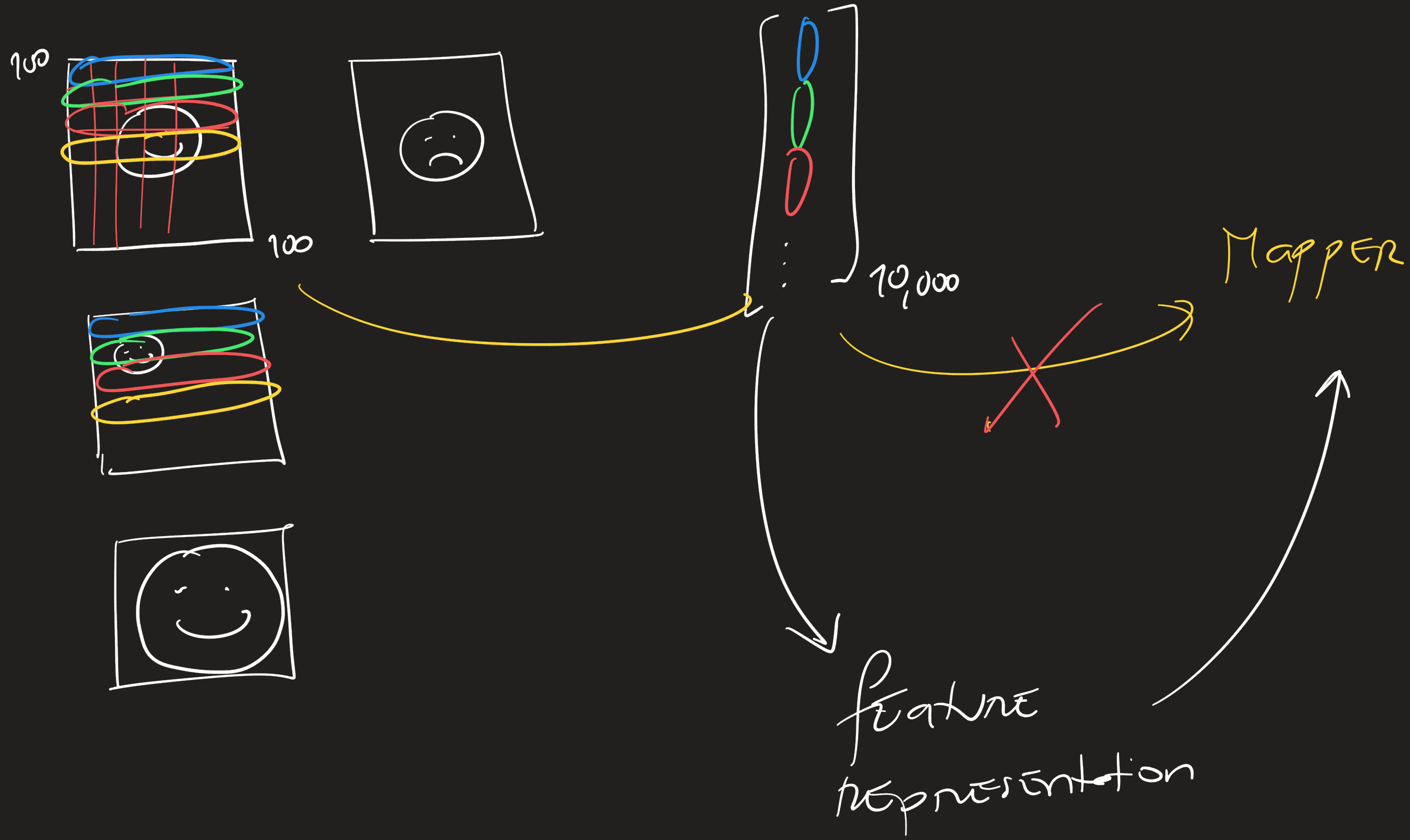

100
100
10,000
Mapper
feature
representation

Spr'24 — Q.5

$$P(x|\mu) = \mu^{x} (1-\mu)^{1-x}$$ — data likelihood

prior

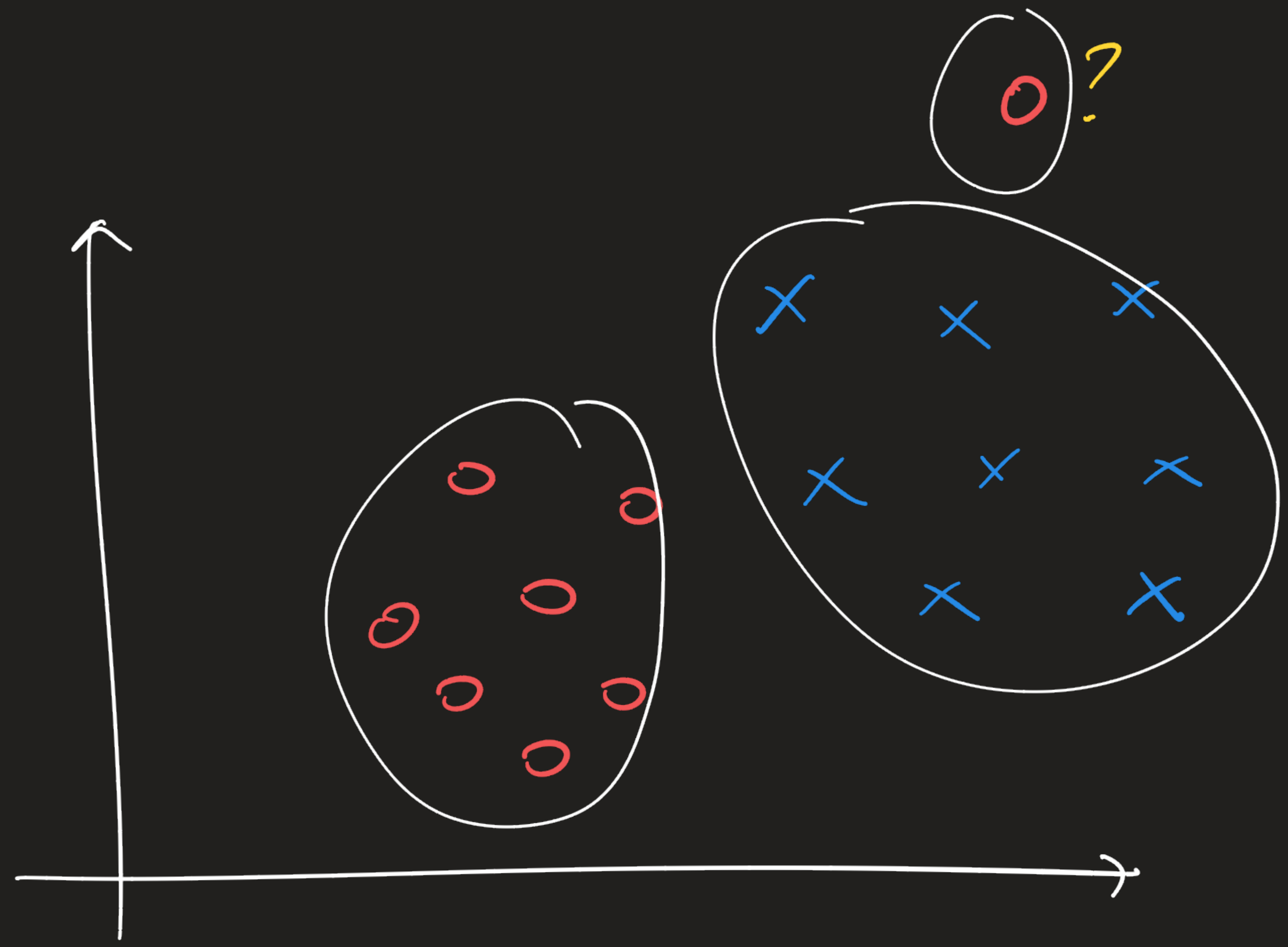
$$P(\mu) = 0.5 \cdot \text{Beta}(\mu|\alpha_1,\beta_1) + 0.5 \cdot \text{Beta}(\mu|\alpha_2,\beta_2)$$

where

$$\text{Beta}(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot x^{a-1} \cdot (1-x)^{b-1}$$

Q. : Write down $\mathcal{L}^{\circ}$, in a simplified form.

(1) $$\mathcal{L}^{\circ} = \left[\prod_{i=1}^{N} P(x_i|\mu)\right] \cdot \left[\frac{1}{2}\text{Beta}(\mu|\alpha_1,\beta_1) + \frac{1}{2}\text{Beta}(\mu|\alpha_2,\beta_2)\right]$$



Censored data

$$x_1, \dots, x_m, \underbrace{x_{m+1}, \dots, x_N}_{\substack{\text{censored samples} \\ \text{(capped at } 100^\circ)}}$$

$$\mathcal{L}^o = \prod_{i=1}^{m} P(x_i | \theta) \cdot \prod_{j=m+1}^{N} \int_{100}^{+\infty} P(x_j | \theta) \cdot dx_j$$

intractable density fct.

Hidden variables : $z_j \equiv$ true value for censored samples, $j = m+1, \dots, N$. $\in \mathbb{R}$

$$\mathcal{L}^c = \prod_{i=1}^{m} P(x_i | \theta) \cdot \prod_{j=m+1}^{N} P(z_j | \theta)$$