

**Answer:**

$$\begin{aligned}\theta_{\text{MAP}} &= \operatorname{argmax}_{\theta} p(\theta|x, y) = \operatorname{argmax}_{\theta} \frac{p(\theta, x, y)}{p(x, y)} = \operatorname{argmax}_{\theta} \frac{p(\theta, x, y)}{p(x, \theta)} \frac{p(x, \theta)}{p(x, y)} \\ &= \operatorname{argmax}_{\theta} p(y|x, \theta) \frac{p(x, \theta)}{p(x)} \frac{p(x)}{p(x, y)} = \operatorname{argmax}_{\theta} p(y|x, \theta)p(x)\end{aligned}$$

Where at the last step we have used the assumption that  $p(\theta) = p(\theta|x)$  (i.e.  $\theta$  is independent from  $x$  (if no knowledge of  $y$ )), and that we can drop the terms with no  $\theta$  since we're argmaxing of  $\theta$