

© Training set: $\{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^m$

$$P(y^{(i)} | \mathbf{x}^{(i)}; \theta) = \frac{1}{y!} \exp((\theta^T \mathbf{x}^{(i)}) y^{(i)} - e^{\theta^T \mathbf{x}^{(i)}})$$

$$l_i(\theta) = \log(P(y^{(i)} | \mathbf{x}^{(i)}; \theta))$$

$$= \log\left(\frac{1}{y!}\right) + (\theta^T \mathbf{x}^{(i)}) y^{(i)} - e^{\theta^T \mathbf{x}^{(i)}}$$

$$\nabla_{\theta} l_i(\theta) = (\mathbf{x}^{(i)}) y^{(i)} - e^{\theta^T \mathbf{x}^{(i)}} \mathbf{x}^{(i)}$$

$$= (y^{(i)} - e^{\theta^T \mathbf{x}^{(i)}}) \mathbf{x}^{(i)}$$

$$\Rightarrow \theta := \theta - \alpha \nabla_{\theta} l_i(\theta)$$

$$\Rightarrow \theta := \theta - \alpha (y^{(i)} - e^{\theta^T \mathbf{x}^{(i)}}) \mathbf{x}^{(i)}$$