

$$J(\theta) = - \sum_i \hat{y}_i \ln(\hat{y}_i)$$