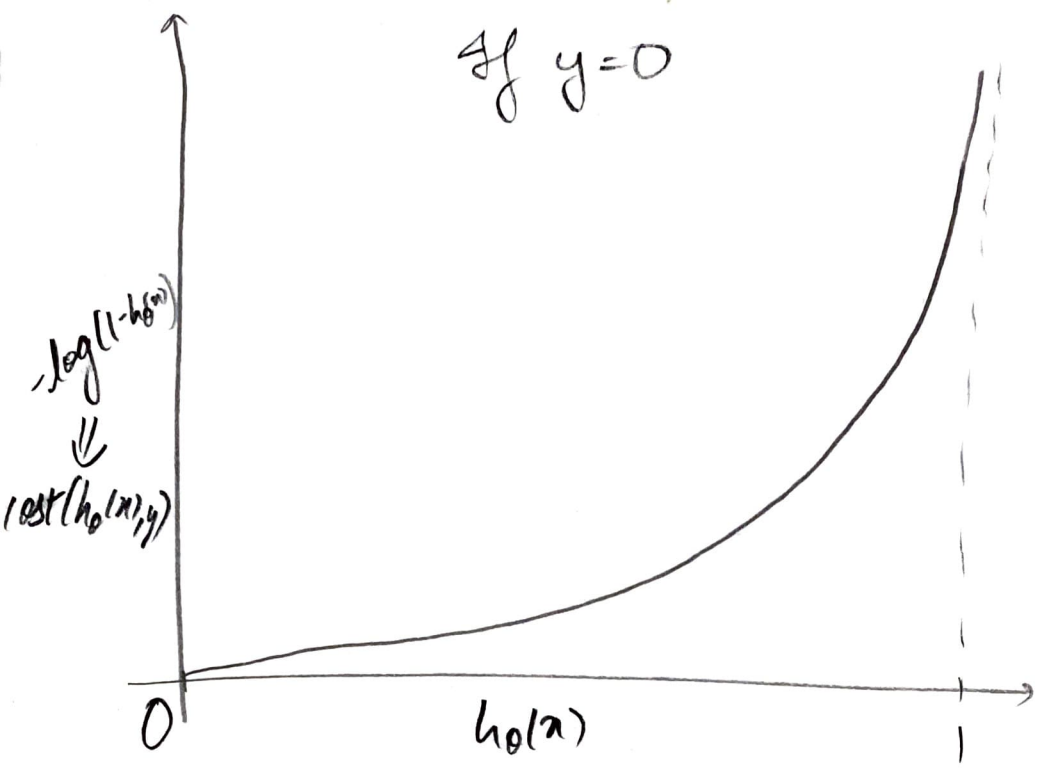

Cost = 0 if $y = 0$, $h_\theta(x) = 0$

But as, $h_\theta(x) \to 1$

$\log y \to \infty$

Captures intuition that if $h_\theta(x) = 0$, (i.e. predict $P(y=0|x;\theta) = 1$), but $y = 0$, then we'll penalize learning algorithm by a very large cost