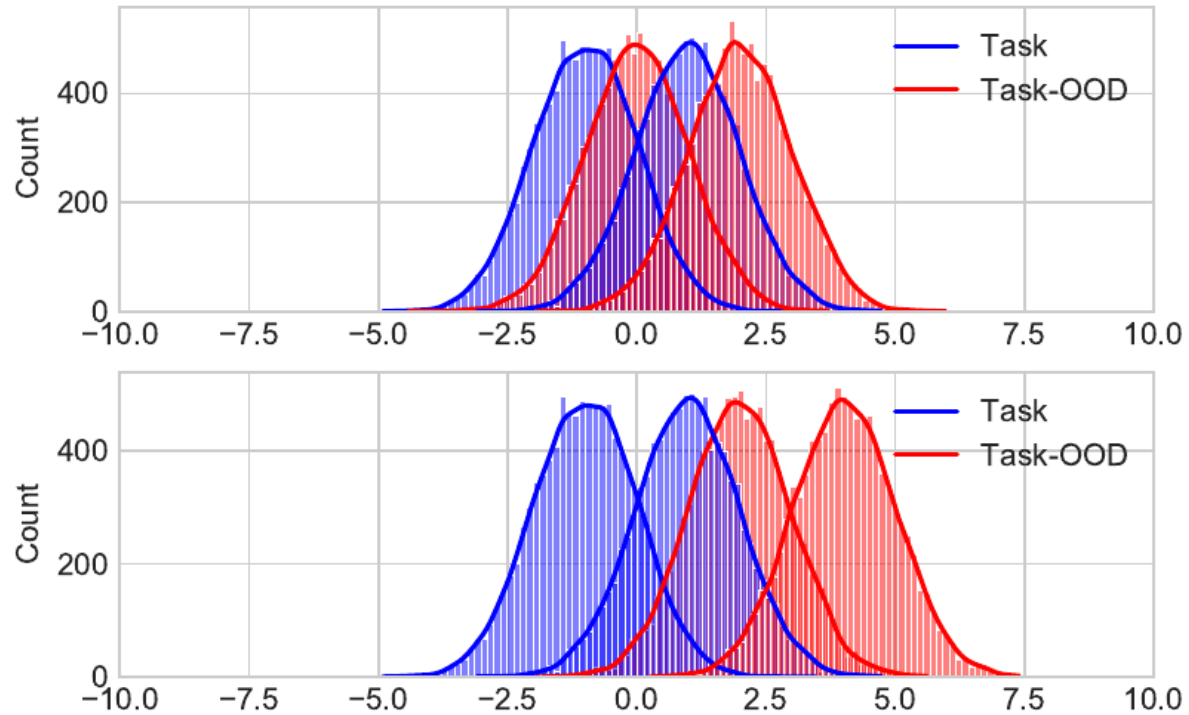


Gaussian Task

We consider a family of tasks

$$P(x, y) = \delta(y = 1)X_1 + \delta(y = -1)X_{-1}$$

where X_1 and X_2 are Gaussians with means $\pm\mu$



Error of Hypothesis

Consider random variable \bar{h} . If $\bar{h} = h$, then

$$e_t(h) = \frac{1}{2} ((1 - \Phi(h + \mu)) + \Phi(h - \mu)) .$$

If we assume $\bar{h} \sim \mathcal{N}(\bar{\mu}, \bar{\sigma})$

$$\mathbb{E}[e_t(\bar{h})] = \frac{1}{2} - \frac{1}{2} \Phi \left(\frac{\bar{\mu} + \mu}{\sqrt{1 + \bar{\sigma}^2}} \right) + \frac{1}{2} \Phi \left(\frac{\bar{\mu} - \mu}{\sqrt{1 + \bar{\sigma}^2}} \right)$$

LDA - Single-head model

Consider the samples to be weighted as

$$S = \alpha S_t + (1 - \alpha) S_{ood}$$

For LDA

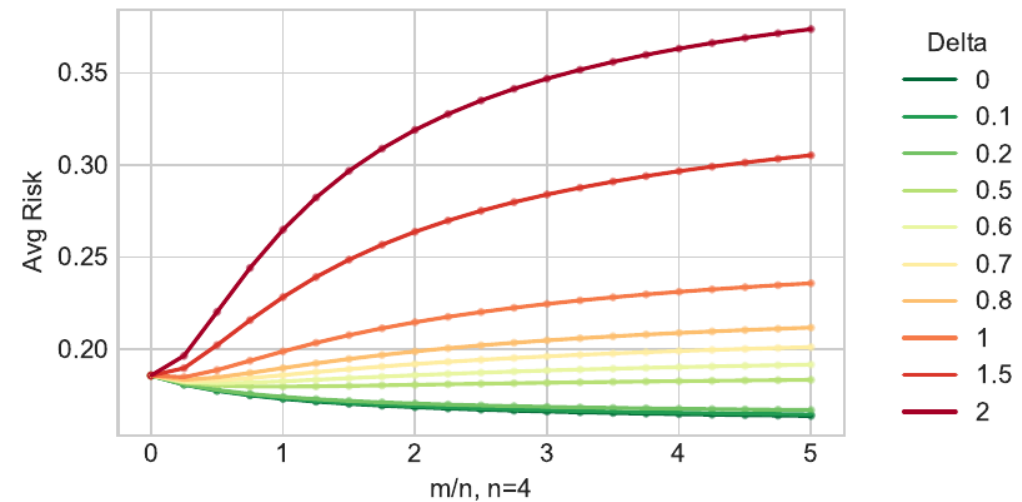
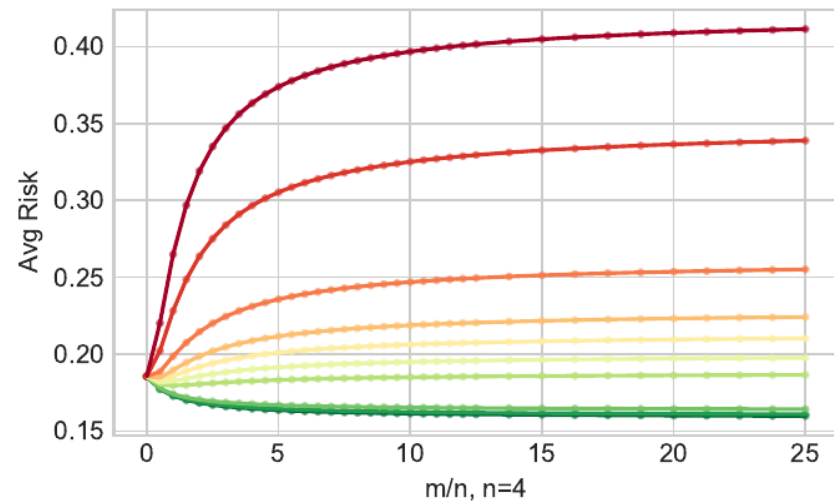
$$\bar{\mu} = \frac{(1 - \alpha)m\Delta}{\alpha n + (1 - \alpha)m}$$

and

$$\bar{\sigma}^2 = \frac{(1 - \alpha)^2 m + \alpha^2 n}{(\alpha n + (1 - \alpha)m)^2}$$

LDA with $\alpha = 0.5$

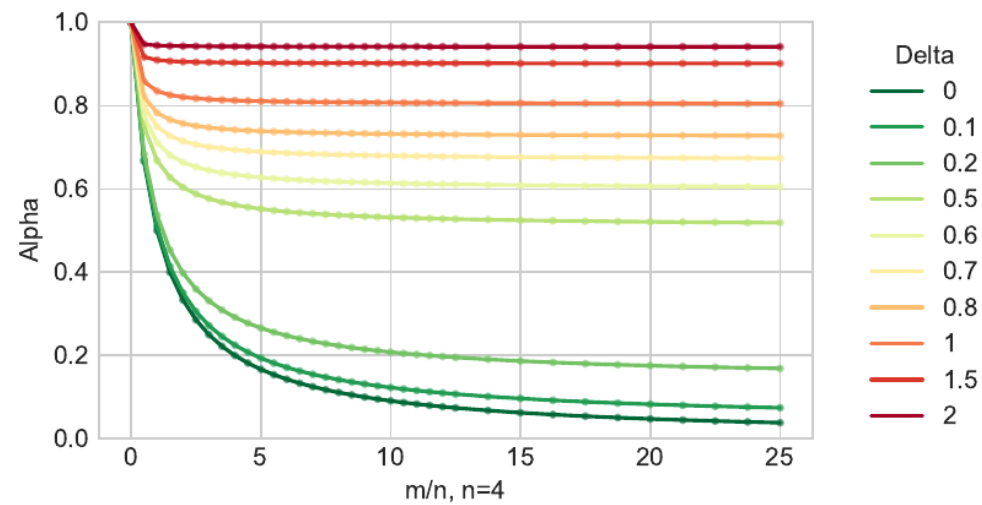
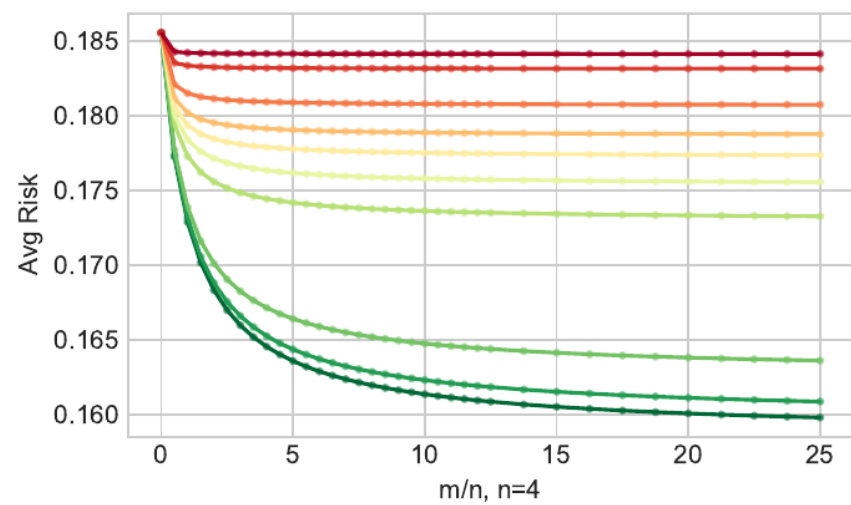
When both datasets are naively combined ($\alpha = 0.5$)



the loss increases/decreases depending on Δ .

LDA with optimized α

However if we optimize α , the loss is always better



Note on optimizing α

α is only usable if we can separate out samples in S_{ood} and S_t . Otherwise, we are forced to use $\alpha = 0.5$.