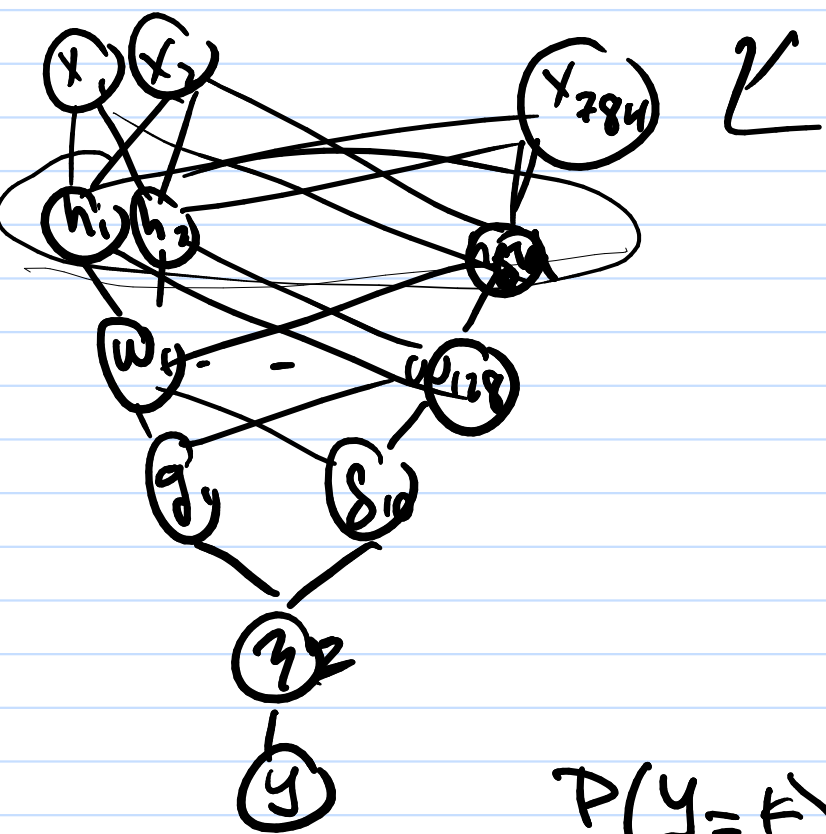
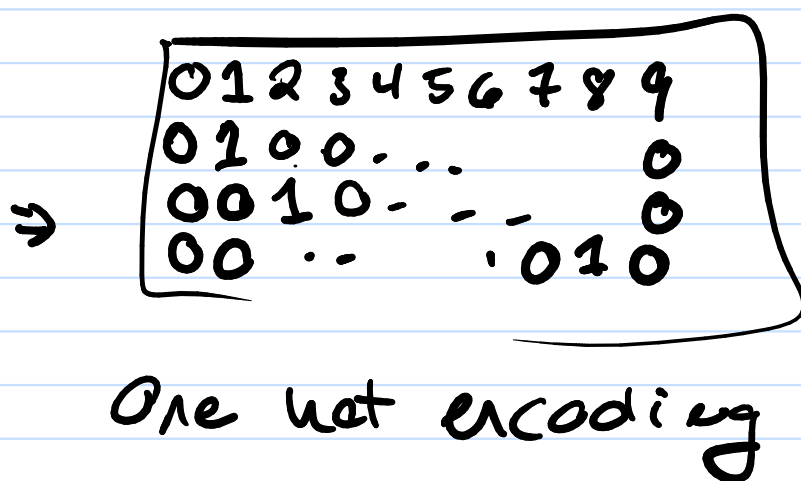
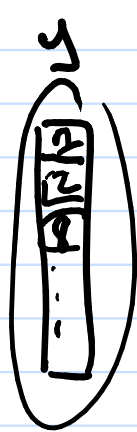
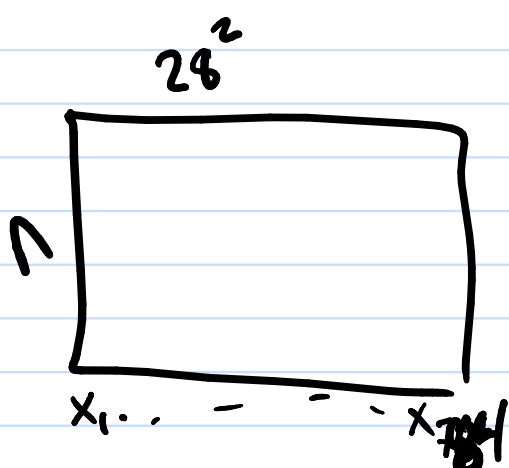
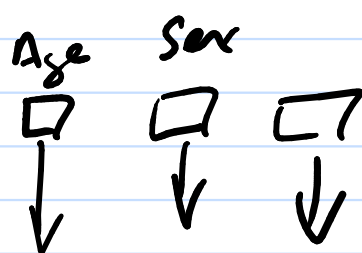
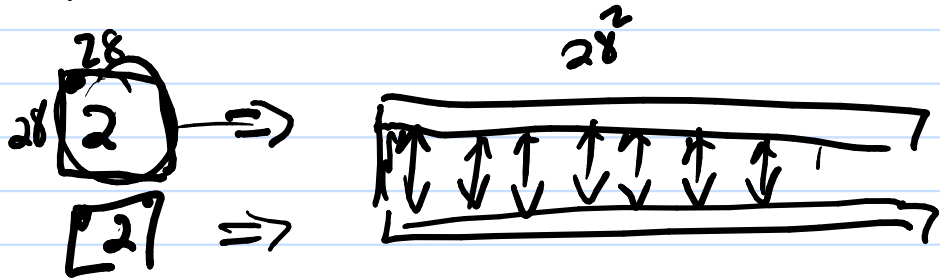


MNIST



$$\frac{e^z}{1+e^z} = P(y=H)$$

$$\frac{1}{1+e^z} = P(y=T)$$

$$P(y=k) = \frac{e^{\mu_k}}{\sum_{k=1}^{10} e^{\mu_k}}$$

$$\sum_k P(y=k) = 1 = \sum_{k=1}^{10} \frac{e^{\mu_k}}{\sum_{k=1}^{10} e^{\mu_k}} = \frac{\sum_{k=1}^{10} e^{\mu_k}}{\sum_{k=1}^{10} e^{\mu_k}} = 1$$

Regularization, Bagging, Random dropout

Tuning param selection:

- Data splitting -
- Cross validation -
- Bootstrapping strategies

