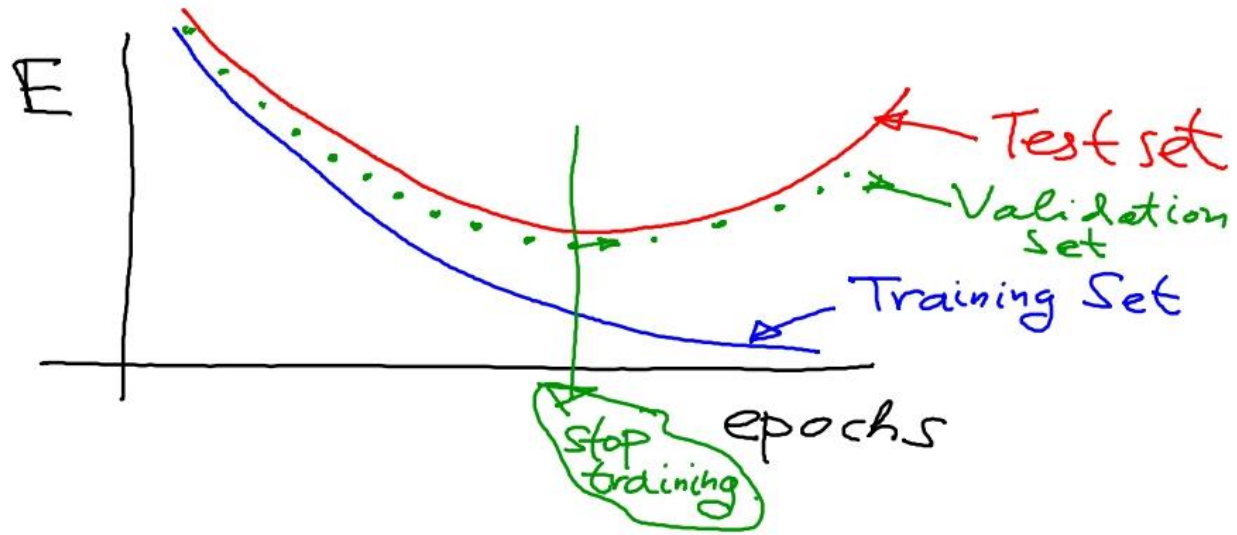
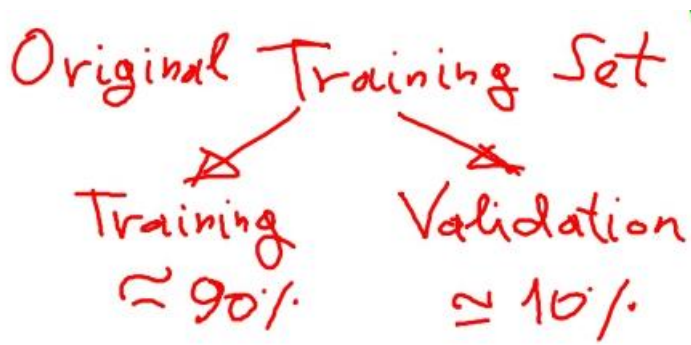
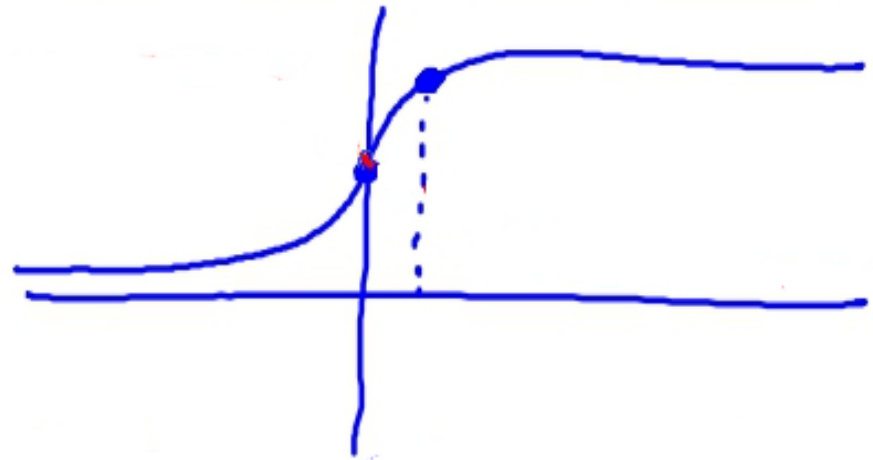

E
Test set
Validation set
Training Set
epochs
Stop training

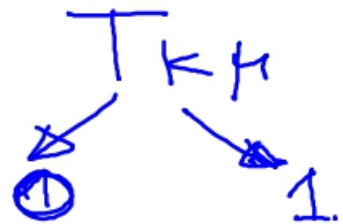

Original Training Set
Training
≃ 90%
Validation
≃ 10%

Cross entropy cost function

$y_{k\mu}^{(R)}$ Probability that my output corresponds to 1.

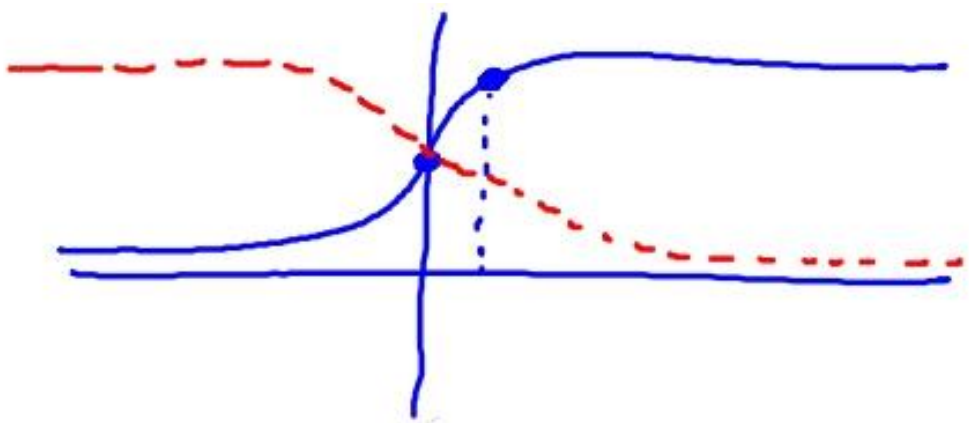


Probability that $y_{k\mu}^{(R)}$ matches its target



(P) $T_{k\mu}=1 \rightarrow y_{k\mu}^{(R)}$

$T_{k\mu}=0 \rightarrow 1-y_{k\mu}^{(R)}$



$$P = y_{k\mu}^{(R)\,T_{k\mu}} \left(1-y_{k\mu}^{(R)}\right)^{1-T_{k\mu}}$$

Likelihood: $$\mathcal{L} = \prod_{pk} y_{kp}^{(R)\,T_{kp}} \left(1 - y_{kp}^{(R)}\right)^{1-T_{kp}}$$

Log likelihood: $$\sum_{pk} \left[T_{kp} \ln y_{kp}^{(R)} + (1 - T_{kp}) \ln\left(1 - y_{kp}^{(R)}\right) \right]$$

Cost function: $$E = -\sum_{pk} \left[T_{kp} \ln y_{kp}^{(R)} + (1 - T_{kp}) \ln\left(1 - y_{kp}^{(R)}\right) \right]$$

$$\frac{\partial E'}{\partial y_{k\mu}^{(R)}} = \frac{T_{k\mu}}{y_{k\mu}^{(R)}} - \frac{1-T_{k\mu}}{1-y_{k\mu}^{(R)}} = \frac{T_{k\mu}-y_{k\mu}^{(R)}}{y_{k\mu}^{(R)}\left(1-y_{k\mu}^{(R)}\right)}$$

1 layered network:

Logistic regression algorithm

Purely classification algorithm