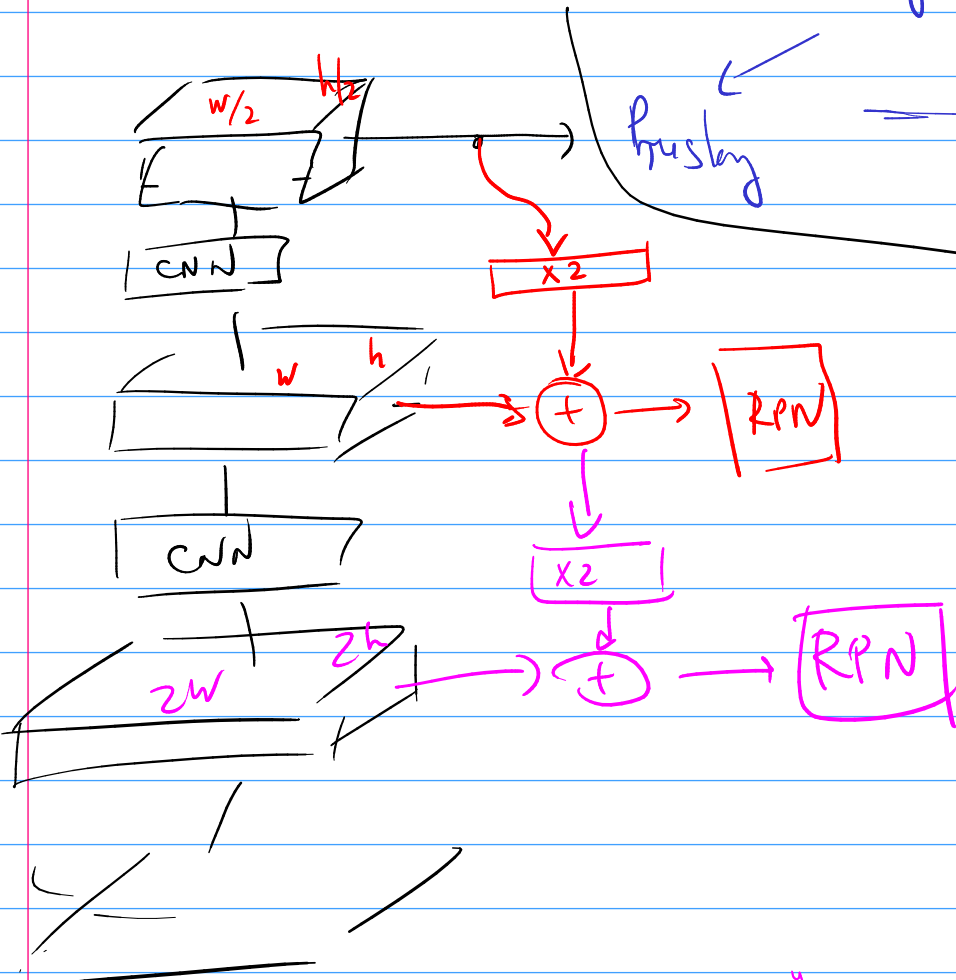
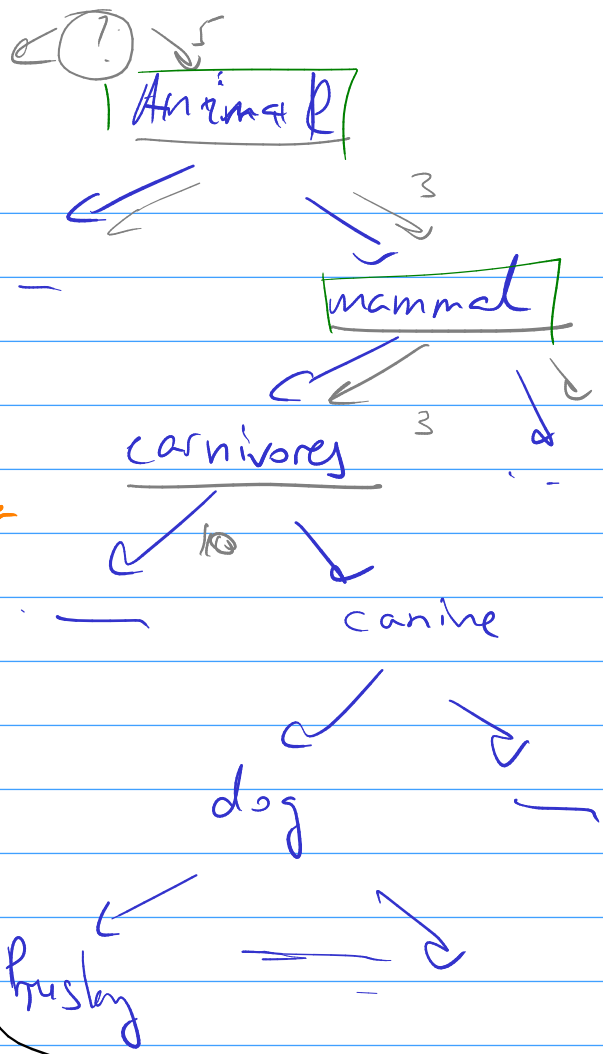
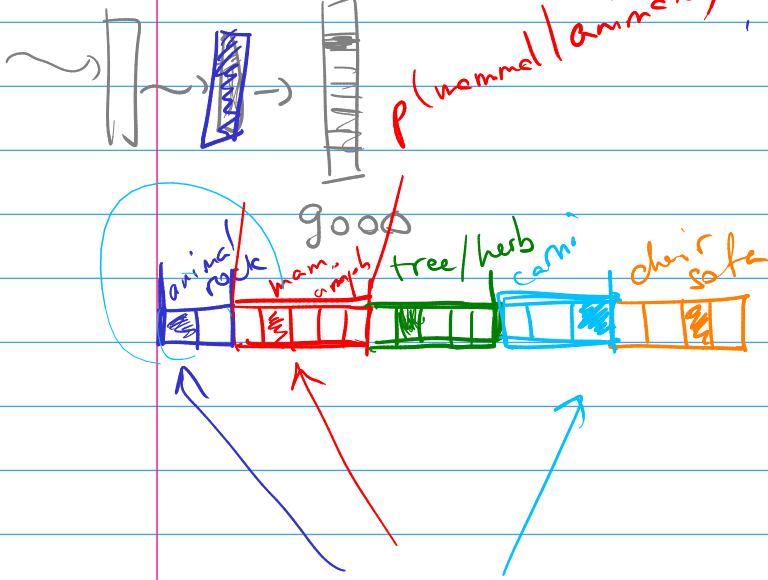


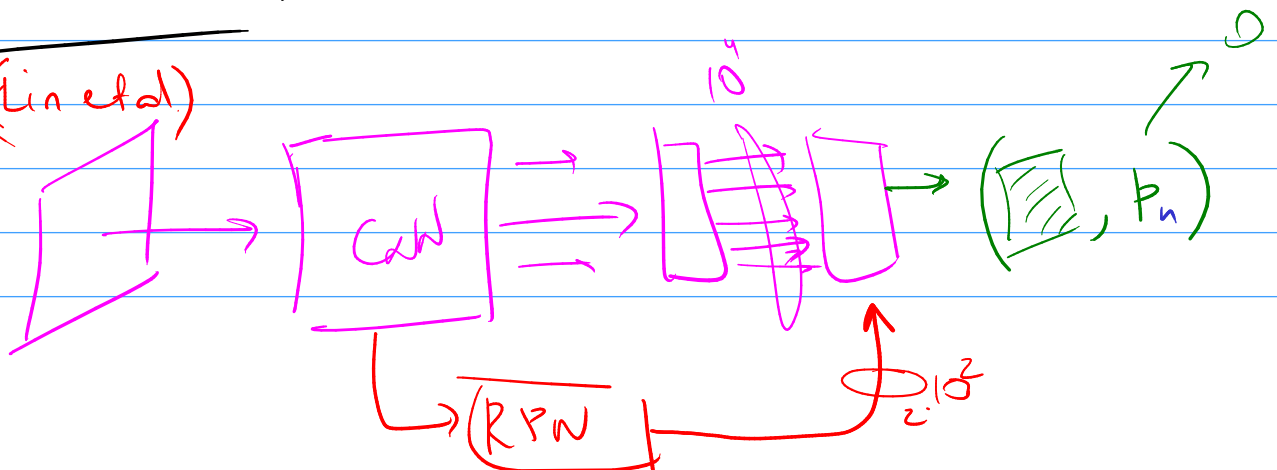
① Hierarchical softmax

goods

softmax



(Lin et al)



$$BCE(p_n, t_n) = -t_n \log p_n - (1 - t_n) \log(1 - p_n) \xrightarrow{p_n \text{ min}}$$

$$t_n = 0 \quad p_n = 1/5$$

$$BCE = -\log 1/5 \approx 0.1 \times 10000$$

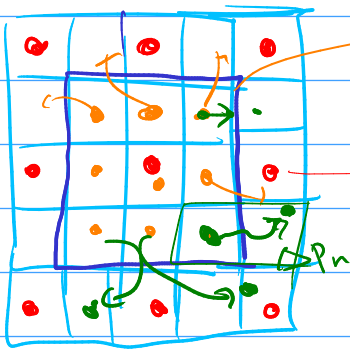
$$p_n = 4/5$$

$$BCE = -\log 1/5 \approx 0.7$$

$$\hat{p}_n = \begin{cases} p_n, & t_n = 1 \\ 1 - p_n, & t_n = 0 \end{cases}$$

$$BCE = -\log \hat{p}_n$$

$$\text{FocalLoss}(\hat{p}_n) = -(1 - \hat{p}_n)^\gamma \log \hat{p}_n$$



$$\bar{X}(p_n)^T \bar{w} \rightarrow h \rightarrow \text{---}$$

$$\bar{X}(p_n)^T \bar{w} \rightarrow h \rightarrow \text{---}$$

$$\bar{X}(p_n + \Delta p_n)^T \bar{w} \rightarrow h \rightarrow \text{---}$$

